CRF

Access DB# 63874

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: Lauren Wells Examiner #: 75202 Date: 4/2/02
Art Unit: 1617 Phone Number 30[illegible] Serial Number: 09/857276
Mail Box and Bldg/Room Location: 3D12 Results Format Preferred (circle): PAPER DISK E-MAIL
2B19

If more than one search is submitted, please prioritize searches in order of need. mej

****

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

_____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

SEQ ID NO: 1032

→ therapeutic peptide

Point of Contact:
Toby Port
Technical Info. Specialist
CM1 6A04
703-308-3534

prot 1032

****

STAFF USE ONLY

Searcher: _____
Searcher Phone #: _____
Searcher Location: _____
Date Searcher Picked Up: 4/5
Date Completed: 4/8
Searcher Prep & Review Time: 10
Clerical Prep Time: _____
Online Time: 10

Type of Search

NA Sequence (#) _____
AA Sequence (#) 1
Structure (#) _____
Bibliographic _____
Litigation _____
Fulltext _____
Patent Family _____
Other _____

Vendors and cost where applicable

STN _____
Dialog _____
Questel/Orbit _____
Dr.Link _____
Lexis/Nexis _____
Sequence Systems CS
WWW/Internet _____
Other (specify) _____

PTO-1590 (8-01)

```
  Query Match           100.0%;  Score 138;  DB 2;  Length 222;
  Best Local Similarity  100.0%;  Pred. No. 6.9e-13;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 NQGRHFCGGALIHARFVMTAASCFQ 44


RESULT  12
US-08-944-483-32
; Sequence 32, Application US/08944483
; Patent No. 6232456
;  GENERAL INFORMATION:
;    APPLICANT:  COHEN, MAURICE
;    APPLICANT:  COLPITTS, TRACEY L.
;    APPLICANT:  FRIEDMAN, PAULA N.
;    APPLICANT:  GRANADOS, EDWARD N.
;    APPLICANT:  KLASS, MICHAEL R.
;    APPLICANT:  RUSSELL, JOHN C.
;    APPLICANT:  STEWART, KENT D.
;    APPLICANT:  STROUPE, STEVEN D.
;    TITLE OF INVENTION:  NOVEL SERINE PROTEASE REAGENTS
;    TITLE OF INVENTION:  AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
;    TITLE OF INVENTION:  OF THE PROSTATE
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,483
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6183.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  225 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6232456e
US-08-944-483-32


  Query Match           100.0%;  Score 138;  DB 4;  Length 225;
  Best Local Similarity  100.0%;  Pred. No. 7e-13;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 NQGRHFCGGALIHARFVMTAASCFQ 44


RESULT  13
US-09-258-934-38
; Sequence 38, Application US/09258934
; Patent No. 6107460
; GENERAL INFORMATION:
;  APPLICANT: Anne Pereira
;  TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF USE THEREOF
;  FILE REFERENCE: 5864.002
;  CURRENT APPLICATION NUMBER: US/09/258,934
;  CURRENT FILING DATE: 1999-03-01
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: WordPerfect 5.1 (saved in ASCII format)  *Software*
; SEQ ID NO 38
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-258-934-38


  Query Match           99.3%;  Score 137;  DB 3;  Length 25;
  Best Local Similarity  96.0%;  Pred. No. 1.2e-13;
  Matches   24;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||:||||||||
Db       1 NQGRHFCGGALIHARFIMTAASCFQ 25


RESULT  14
US-09-258-934-36
; Sequence 36, Application US/09258934
; Patent No. 6107460
; GENERAL INFORMATION:
;  APPLICANT: Anne Pereira
;  TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF USE THEREOF
;  FILE REFERENCE: 5864.002
;  CURRENT APPLICATION NUMBER: US/09/258,934
;  CURRENT FILING DATE: 1999-03-01
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: WordPerfect 5.1 (saved in ASCII format)  *Software*
; SEQ ID NO 36
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: homo sapien
US-09-258-934-36


  Query Match           97.8%;  Score 135;  DB 3;  Length 25;
  Best Local Similarity  96.0%;  Pred. No. 2.2e-13;
  Matches   24;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||:|||
Db       1 NQGRHFCGGALIHARFVMTAATCFQ 25


RESULT  15
US-09-258-934-37
; Sequence 37, Application US/09258934
; Patent No. 6107460
; GENERAL INFORMATION:
;  APPLICANT: Anne Pereira
;  TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF USE THEREOF
;  FILE REFERENCE: 5864.002
;  CURRENT APPLICATION NUMBER: US/09/258,934
;  CURRENT FILING DATE: 1999-03-01
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: WordPerfect 5.1 (saved in ASCII format)  *Software*
; SEQ ID NO 37
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: homo sapien
```

```
;      OPERATING SYSTEM:  MS-DOS 5.00
;      SOFTWARE:  WordPerfect 5.1 (saved in ASCII format)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/260,373
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/482,328
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,931
;      FILING DATE:  October 30, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/855,417
;      FILING DATE:  March 18, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/543,151
;      FILING DATE:  June 25, 1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/375,739
;      FILING DATE:  July 5, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corbett, Christopher W., Ph.D.
;      REGISTRATION NUMBER:  36,105
;      REFERENCE/DOCKET NUMBER:  5820.360
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  405-478-5344
;      TELEFAX:  405-478-5349
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:  20-44aa of mature CAP37 protein
US-09-260-373-1


  Query Match             100.0%;  Score 138;  DB 3;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   7
US-09-258-934-1
; Sequence 1, Application US/09258934
; Patent No. 6107460
; GENERAL INFORMATION:
;  APPLICANT: Anne Pereira
;  TITLE OF INVENTION: ANTIMICROBIAL PEPTIDES AND METHODS OF USE THEREOF
;  FILE REFERENCE: 5864.002
;  CURRENT APPLICATION NUMBER: US/09/258,934
;  CURRENT FILING DATE: 1999-03-01
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: WordPerfect 5.1 (saved in ASCII format)  *Software*
; SEQ ID NO 1
;    LENGTH: 25
;    TYPE: PRT
; . ORGANISM: homo sapien
US-09-258-934-1


  Query Match             100.0%;  Score 138;  DB 3;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   8
US-08-925-708-1
; Sequence 1, Application US/08925708
; Patent No. 5939390
;  GENERAL INFORMATION:
;    APPLICANT:  Flodgaard, Hans
;    APPLICANT:  Rasmussen, Poul Baad
;    TITLE OF INVENTION:  A Pharmaceutical Composition
;  * NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  No. 5939390o No. 5939390disk of No. 5939390th America, Inc.
;      STREET:  405 Lexington Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10174
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/925,708
;      FILING DATE:  9-SEPT-1997
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Agris, Cheryl H.
;      REGISTRATION NUMBER:  34,086
;      REFERENCE/DOCKET NUMBER:  4319.204-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-867-0123
;      TELEFAX:  212-878-9655
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  221 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    ORIGINAL SOURCE:
;      ORGANISM:  human
US-08-925-708-1


  Query Match             100.0%;  Score 138;  DB 2;  Length 221;
  Best Local Similarity   100.0%;  Pred. No. 6.9e-13;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 NQGRHFCGGALIHARFVMTAASCFQ 44


RESULT   9
US-07-969-931-9
; Sequence 9, Application US/07969931
; Patent No. 5458874
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne
;    APPLICANT:  Spitznagel, John K.
;    TITLE OF INVENTION:  Chemotactic, Antibiotic and
;    TITLE OF INVENTION:  Lipopolysaccharide-Binding Peptide Fragments of CAP37
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Needle & Rosenberg, P.C.
;      STREET:  133 Carnegie Way N.W., Suite 400
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
```

```
  Query Match           100.0%;  Score 138;  DB 1;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   2
US-07-855-417A-8
; Sequence 8, Application US/07855417A
; Patent No. 5484885
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne
;    APPLICANT:  Spitznagel, John K.
;    TITLE OF INVENTION:  Chemotactic, Antibiotic and
;    TITLE OF INVENTION:  Lipopolysaccharide-Binding Peptide Fragments of CAP37
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Needle & Rosenberg, P.C.
;      STREET:  133 Carnegie Way N.W., Suite 400
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/855,417A
;      FILING DATE:  18-MAR-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/543,151
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/375,739
;      FILING DATE:  05-JUL-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Severson, Mary L.
;      REGISTRATION NUMBER:  34,927
;      REFERENCE/DOCKET NUMBER:  0510.024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:  20-44aa of mature CAP37 protein
US-07-855-417A-8


  Query Match           100.0%;  Score 138;  DB 1;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   3
US-08-235-399-1
; Sequence 1, Application US/08235399
; Patent No. 5607916
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne; Brackett, Daniel J.; and
;    APPLICANT:  Lerner, Megan, R.
;    TITLE OF INVENTION:  Method and Composition for The Treatment
;    TITLE OF INVENTION:  of Septic Shock
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christopher W. Corbett, Ph.D.
;      ADDRESSEE:  Dunlap, Codding & Lee, P.C.
;      STREET:  9400 No. 5607916th Broadway, Suite 420
;      CITY:  Oklahoma City
;      STATE:  Oklahoma
;      COUNTRY:  United States of America
;      ZIP:  73114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5, 1.44 Mb High Density Diskette
;      COMPUTER:  IBM AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.00
;      SOFTWARE:  WordPerfect 5.1 (saved in ASCII format)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/235,399
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,931
;      FILING DATE:  October 30, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/855,417
;      FILING DATE:  March 18, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/543,151
;      FILING DATE:  June 25, 1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/375,739
;      FILING DATE:  July 5, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Christopher W. Corbett, Ph.D.
;      REGISTRATION NUMBER:  36,109
;      REFERENCE/DOCKET NUMBER:  5820.291
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  405-478-5344
;      TELEFAX:  405-478-5349
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:  20-44aa of mature CAP37 protein
US-08-235-399-1


  Query Match           100.0%;  Score 138;  DB 1;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   4
US-08-482-328-1
; Sequence 1, Application US/08482328
; Patent No. 5627262
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne
;    TITLE OF INVENTION:  Method and Composition for The Treatment
;    TITLE OF INVENTION:  of Septic Shock
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christopher W. Corbett, Ph.D.
;      ADDRESSEE:  Dunlap & Codding, P.C.
```

```
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AE003792; AAF57469.1; -.
DR   HSSP; P00763; 1DPO.
DR   FlyBase; FBgn0034507; CG11192.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   269 AA;  28827 MW;  9797839B70ECD318 CRC64;


  Query Match           60.1%;  Score 83;  DB 5;  Length 269;
  Best Local Similarity   62.5%;  Pred. No. 1.3e-05;
  Matches   15;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCFQ 25
           |||| ||||:|    |:||| ||:
Db      38 QGRHICGGAIIGIDTVLTAAHCFE 61


RESULT  13
Q13208
ID   Q13208      PRELIMINARY;      PRT;   567 AA.
AC   Q13208;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   HEPATOCYTE GROWTH FACTOR-LIKE PROTEIN HOMOLOG (D1F15S1A).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Degen S.J.F., McDowell S.A., Waltz S.E., Gould F., Stuart L.A.,
RA   Carritt B.;
RL   Submitted (MAY-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
DR   EMBL; U28054; AAC63092.1; -.
DR   HSSP; P00747; 2PK4.
DR   MEROPS; S01.977; -.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   PROSITE; PS00021; KRINGLE_1; 2.
DR   PROSITE; PS50070; KRINGLE_2; 4.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   567 AA;  64116 MW;  3FC38B07F1645810 CRC64;


  Query Match           60.1%;  Score 83;  DB 4;  Length 567;
  Best Local Similarity   52.2%;  Pred. No. 2.7e-05;
  Matches   12;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCF 24
           ||:|||||:|:  ::::||  ||
Db     477 QGQHFCGGSLVKEQWILTAQCF 499


RESULT  14
Q9H1V4
ID   Q9H1V4      PRELIMINARY;      PRT;   648 AA.
AC   Q9H1V4;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DJ1182A14.3 (SIMILAR TO MST1 (MACROPHAGE STIMULATING 1 (HEPATOCYTE
DE   GROWTH FACTOR-LIKE))).
GN   DJ1182A14.3.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Bird C.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AL137798; CAC17639.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 1.
DR   PROSITE; PS50070; KRINGLE_2; 4.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   648 AA;  72781 MW;  4CE077057350E463 CRC64;


  Query Match           60.1%;  Score 83;  DB 4;  Length 648;
  Best Local Similarity   52.2%;  Pred. No. 3.1e-05;
  Matches   12;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCF 24
           ||:|||||:|:  ::::||  ||
Db     454 QGQHFCGGSLVKEQWILTAQCF 476


RESULT  15
Q14870
ID   Q14870      PRELIMINARY;      PRT;   711 AA.
AC   Q14870;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MACROPHAGE-STIMULATING PROTEIN PRECURSOR.
GN   MST1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE=93340141; PubMed=8393443;
RA   Yoshimura T., Yuhki N., Wang M.H., Skeel A., Leonard E.J.;
RT   "Cloning, sequencing, and expression of human macrophage stimulating
RT   protein (MSP, MST1) confirms MSP as a member of the family of kringle
RT   proteins and locates the MSP gene on chromosome 3.";
RL   J. Biol. Chem. 268:15461-15468(1993).
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; L11924; AAA59872.1; -.
DR   HSSP; P00747; 2PK4.
```

```
FT   NON_TER     117    117
SQ   SEQUENCE   117 AA;  13038 MW;  37052EE97322EC56 CRC64;


  Query Match             61.6%;  Score 85;  DB 13;  Length 117;
  Best Local Similarity   52.0%;  Pred. No. 2.9e-06;
  Matches   13;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:|::| :|::|| |::
Db      38 NSGYHFCGGSLVNAEWVVSAAHCYK 62


RESULT   8
O42608
ID   O42608      PRELIMINARY;      PRT;   247 AA.
AC   O42608;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   TRYPSINOGEN A3 PRECURSOR.
GN   TRYPA3.
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Roach J.C.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=ANTERIOR INTESTINE;
RA   Roach J.C.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF011899; AAB69655.1; -.
DR   EMBL; AF011352; AAB65411.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL       1     15       POTENTIAL.
FT   CHAIN       16    247       TRYPSIN A3.
SQ   SEQUENCE   247 AA;  26295 MW;  BECBD3069A071DCB CRC64;


  Query Match             61.6%;  Score 85;  DB 13;  Length 247;
  Best Local Similarity   56.0%;  Pred. No. 5.9e-06;
  Matches   14;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |:|
Db      42 NIGYHFCGGSLINSQWVVSAAHCYQ 66


RESULT   9
O42158
ID   O42158      PRELIMINARY;      PRT;   247 AA.
AC   O42158;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   TRYPSINOGEN A2 PRECURSOR.
GN   TRYPA2.
OS   Petromyzon marinus (Sea lamprey).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC   Petromyzontiformes; Petromyzontidae; Petromyzon.
OX   NCBI_TaxID=7757;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Roach J.C.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF011898; AAB69654.1; -.
DR   HSSP; P00760; 1AQ7.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL       1     15       POTENTIAL.
FT   CHAIN       16    247       TRYPSIN A2.
SQ   SEQUENCE   247 AA;  26309 MW;  AD73E88531970324 CRC64;


  Query Match             61.6%;  Score 85;  DB 13;  Length 247;
  Best Local Similarity   56.0%;  Pred. No. 5.9e-06;
  Matches   14;  Conservative    8;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |:|
Db      42 NIGYHFCGGSLINSQWVVSAAHCYQ 66


RESULT  10
Q9D7Y7
ID   Q9D7Y7      PRELIMINARY;      PRT;   247 AA.
AC   Q9D7Y7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   2210010C04RIK PROTEIN.
GN   2210010C04RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=STOMACH;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AK008695; BAB25837.1; -.
```

```
RESULT   2
Q61096
ID   Q61096      PRELIMINARY;      PRT;   254 AA.
AC   Q61096;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-AUG-1999 (TrEMBLrel. 11, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PROTEINASE-3 (PROTEINASE 3) (PRE-PRO-PROTEASE 3).
GN   PRTN3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SVJ;
RA   Sturrock A., Franklin K.F., Wu S.Q., Hoidal J.R.;
RT   "Characterization and localization of the genes for mouse proteinase-3
RT   and neutrophil elastase.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 2-254 FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE=97362044; PubMed=9211743;
RA   Aveskogh M., Lutzelschwab C., Huang M.R., Hellman L.;
RT   "Characterization of cDNA clones encoding mouse proteinase 3
RT   (myeloblastine) and cathepsin G.";
RL   Immunogenetics 46:181-191(1997).
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF082186; AAC79701.1; -.
DR   EMBL; U43525; AAB67271.1; -.
DR   HSSP; P20160; 1AE5.
DR   MEROPS; S01.134; -.
DR   MGD; MGI:893580; Prtn3.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   254 AA;  27626 MW;  00CEB989A3CB79CA CRC64;


  Query Match           71.0%;  Score 98;  DB 11;  Length 254;
  Best Local Similarity   73.9%;  Pred. No. 4.7e-08;
  Matches   17;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 GRHFCGGALIHARFVMTAASCFQ 25
           | ||||| ||| |||:||| | |
Db      54 GSHFCGGTLIHPRFVLTAAHCLQ 76


RESULT   3
Q9GME1
ID   Q9GME1      PRELIMINARY;      PRT;   258 AA.
AC   Q9GME1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   N-ELASTASE.
OS   Ornithorhynchus anatinus (Duckbill platypus).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.
OX   NCBI_TaxID=9258;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Poorafshar M.M., Hellman L.L.;
RT   "Identification and structural analysis of three serine proteases in a
RT   monotreme, platypus, Ornithorhynchus anatinus.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF275652; AAG00451.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   258 AA;  28165 MW;  2E1752C300A36138 CRC64;


  Query Match           71.0%;  Score 98;  DB 6;  Length 258;
  Best Local Similarity   73.9%;  Pred. No. 4.8e-08;
  Matches   17;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASC 23
           |:| ||||| ||| :|||||| |
Db      43 NRGGHFCGGTLIHQQFVMTAAHC 65


RESULT   4
Q9UQD8
ID   Q9UQD8      PRELIMINARY;      PRT;   200 AA.
AC   Q9UQD8;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PROTEINASE 3 (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Clave E., Molldrem J., Hensel N., Raptis A., Barrett A.J.;
RT   "Donor-recipient polymorphism of the proteinase 3 gene: A potential
RT   target for T-cell alloresponses to myeloid leukemia.";
RL   J. Immunother. 22:1-6(1999).
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF015449; AAD21524.1; -.
DR   EMBL; AF015446; AAD21524.1; JOINED.
DR   EMBL; AF015447; AAD21524.1; JOINED.
DR   EMBL; AF015448; AAD21524.1; JOINED.
DR   HSSP; P24158; 1FUJ.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Hydrolase; Serine protease.
FT   NON_TER     200     200
SQ   SEQUENCE   200 AA;  21609 MW;  13F350B1D0E160A8 CRC64;


  Query Match           67.4%;  Score 93;  DB 4;  Length 200;
  Best Local Similarity   64.0%;  Pred. No. 2.4e-07;
  Matches   16;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | ||||| |||  ||:||| | :
Db      50 NPGSHFCGGTLIHPSFVLTAAHCLR 74


RESULT   5
Q9DBI0
ID   Q9DBI0      PRELIMINARY;      PRT;   799 AA.
AC   Q9DBI0;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
```

```
Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | ||||:|| :::|::|| |::
Db      44 NSGYHFCGGSLISSQWVLSAAHCYK 68
```

Search completed: April 5, 2002, 15:01:12
Job time: 106 sec

```
 Query Match           60.9%;  Score 84;  DB 1;  Length 247;
 Best Local Similarity  56.0%;  Pred. No. 3.9e-06;
 Matches   14;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||: ::|::|| |:|
Db      42 NAGYHFCGGSLINDQWVVSAAHCYQ 66


RESULT  12
TRY2_XENLA
ID   TRY2_XENLA     STANDARD;      PRT;   244 AA.
AC   P70059;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN PRECURSOR (EC 3.4.21.4).
OS   Xenopus laevis (African clawed frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
OC   Xenopodinae; Xenopus.
OX   NCBI_TaxID=8355;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wang K., Lytle L., Gan L., Hood L.E.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U72330; AAB17274.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.258; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL       1     15       BY SIMILARITY.
FT   PROPEP      16     21       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN       22    244       TRYPSIN.
FT   ACT_SITE    61     61       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   105    105       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   198    198       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    28    158       BY SIMILARITY.
FT   DISULFID    46     62       BY SIMILARITY.
FT   DISULFID   130    231       BY SIMILARITY.
FT   DISULFID   137    204       BY SIMILARITY.
FT   DISULFID   169    183       BY SIMILARITY.
FT   DISULFID   194    218       BY SIMILARITY.
FT   SITE       192    192       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   244 AA;  26079 MW;  C63F29CB3300B323 CRC64;


 Query Match           60.1%;  Score 83;  DB 1;  Length 244;
 Best Local Similarity  52.0%;  Pred. No. 5.6e-06;
 Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||:::|::|| |::
Db      40 NAGYHFCGGSLINSQWVVSAAHCYK 64


RESULT  13
TRY3_RAT
ID   TRY3_RAT       STANDARD;      PRT;   247 AA.
AC   P08426;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-AUG-1988 (Rel. 08, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN III, CATIONIC PRECURSOR (EC 3.4.21.4) (PRETRYPSINOGEN III).
GN   TRY3.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87271609; PubMed=3607011;
RA   Fletcher T.S., Alhadeff M., Craik C.S., Largman C.;
RT   "Isolation and characterization of a cDNA encoding rat cationic
RT   trypsinogen.";
RL   Biochemistry 26:3081-3086(1987).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M16624; AAA41985.1; -.
DR   PIR; A27547; A27547.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL       1     15
FT   PROPEP      16     24       ACTIVATION PEPTIDE.
FT   CHAIN       25    247       TRYPSIN III, CATIONIC.
FT   ACT_SITE    64     64       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   108    108       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   201    201       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    31    161       BY SIMILARITY.
FT   DISULFID    49     65       BY SIMILARITY.
FT   DISULFID   133    234       BY SIMILARITY.
FT   DISULFID   140    207       BY SIMILARITY.
FT   DISULFID   172    186       BY SIMILARITY.
FT   DISULFID   197    221       BY SIMILARITY.
FT   SITE       195    195       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   247 AA;  26269 MW;  D74892BAA584E4A8 CRC64;


 Query Match           60.1%;  Score 83;  DB 1;  Length 247;
 Best Local Similarity  52.0%;  Pred. No. 5.7e-06;
 Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
```

```
RX   MEDLINE=66079271; PubMed=5892911;
RA   Kauffman D.L.;
RT   "The disulphide bridges of trypsin.";
RL   J. Mol. Biol. 12:929-932(1965).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: SYNTHESIZED IN THE ACINAR CELLS OF THE
CC       PANCREAS.
CC   -!- PTM: AUTOCATALYTIC CLEAVAGE AFTER LYS-20 LEADS TO BETA-TRYPSIN BY
CC       RELEASING A TERMINAL HEXAPEPTIDE. SUBSEQUENT CLEAVAGE AFTER
CC       LYS-145 LEADS TO ALPHA-TRYPSIN. FURTHER CLEAVAGE AFTER LYS-190
CC       YIELDS PSEUDOTRYPSIN. A CLEAVAGE MAY ALSO OCCUR AFTER ARG-119.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   -!- DATABASE: NAME=Worthington enzyme manual;
CC       WWW="http://www.worthington-biochem.com/manual/T/TRY.html".
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D38507; BAA07516.1; -.
DR   PIR; A00946; TRBOTR.
DR   PDB; 2TGA; 09-APR-85.
DR   PDB; 1TGB; 14-MAR-85.
DR   PDB; 1TGC; 09-APR-85.
DR   PDB; 2TGD; 07-MAY-86.
DR   PDB; 1TGN; 22-OCT-84.
DR   PDB; 2TGP; 14-MAR-85.
DR   PDB; 1TGS; 14-MAR-85.
DR   PDB; 1TGT; 15-OCT-90.
DR   PDB; 2TGT; 09-APR-85.
DR   PDB; 1TLD; 15-OCT-92.
DR   PDB; 2TLD; 15-JUL-92.
DR   PDB; 1TPA; 14-MAR-85.
DR   PDB; 2TPI; 15-JUL-91.
DR   PDB; 3TPI; 14-MAR-85.
DR   PDB; 4TPI; 08-NOV-85.
DR   PDB; 1TPO; 14-MAR-85.
DR   PDB; 1TPP; 16-APR-87.
DR   PDB; 1NTP; 15-APR-91.
DR   PDB; 3PTB; 14-MAR-85.
DR   PDB; 2PTC; 14-MAR-85.
DR   PDB; 2PTN; 15-APR-91.
DR   PDB; 3PTN; 09-APR-85.
DR   PDB; 5PTP; 07-JUL-97.
DR   PDB; 1TAB; 15-JUL-92.
DR   PDB; 1GBT; 31-JAN-94.
DR   PDB; 1PPC; 31-JAN-94.
DR   PDB; 1PPE; 31-JAN-94.
DR   PDB; 1PPH; 31-JAN-94.
DR   PDB; 1SMF; 31-JUL-94.
DR   PDB; 1TNG; 30-NOV-94.
DR   PDB; 1TNH; 30-NOV-94.
DR   PDB; 1TNI; 30-NOV-94.
DR   PDB; 1TNJ; 30-NOV-94.
DR   PDB; 1TNK; 30-NOV-94.
DR   PDB; 1TNL; 30-NOV-94.
DR   PDB; 1BTP; 29-JAN-96.
DR   PDB; 1BTW; 15-OCT-95.
DR   PDB; 1BTX; 15-OCT-95.
DR   PDB; 1BTY; 15-OCT-95.
DR   PDB; 1BTZ; 15-OCT-95.
DR   PDB; 1JRS; 14-OCT-96.
DR   PDB; 1JRT; 14-OCT-96.
DR   PDB; 1MAX; 14-OCT-96.
DR   PDB; 1MAY; 14-OCT-96.
DR   PDB; 1MTS; 20-AUG-97.
DR   PDB; 1MTU; 12-NOV-97.
DR   PDB; 1MTV; 12-NOV-97.
DR   PDB; 1MTW; 12-NOV-97.
DR   PDB; 1TPS; 26-JAN-95.
DR   PDB; 1TYN; 26-JAN-95.
DR   PDB; 1TAW; 24-JUN-97.
DR   PDB; 1AQ7; 25-FEB-98.
DR   PDB; 1TIO; 23-SEP-98.
DR   PDB; 2TIO; 30-SEP-98.
DR   PDB; 1XUF; 16-DEC-98.
DR   PDB; 1XUG; 16-DEC-98.
DR   PDB; 1XUH; 11-NOV-98.
DR   PDB; 1XUI; 11-NOV-98.
DR   PDB; 1XUJ; 11-NOV-98.
DR   PDB; 1XUK; 11-NOV-98.
DR   PDB; 1AUJ; 14-OCT-98.
DR   PDB; 1AZ8; 13-JAN-99.
DR   PDB; 1BJU; 13-JAN-99.
DR   PDB; 1BJV; 13-JAN-99.
DR   PDB; 2BZA; 23-MAR-99.
DR   PDB; 1CE5; 23-MAR-99.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   3D-structure.
FT   NON_TER       1      1
FT   SIGNAL       <1     14
FT   PROPEP        15     20       ACTIVATION PEPTIDE.
FT   CHAIN         21    243       TRYPSIN, CATIONIC.
FT   CHAIN         21    145       ALPHA-TRYPSIN CHAIN 1.
FT   CHAIN        146    243       ALPHA-TRYPSIN CHAIN 2.
FT   CA_BIND       72     82
FT   BINDING      191    192       SUBSTRATE.
FT   BINDING      194    195       SUBSTRATE.
FT   BINDING      197    197       SUBSTRATE.
FT   DISULFID      27    157
FT   DISULFID      45     61
FT   DISULFID     129    230
FT   DISULFID     136    203
FT   DISULFID     168    182
FT   DISULFID     193    217
FT   ACT_SITE      60     60       CHARGE RELAY SYSTEM.
FT   ACT_SITE     104    104       CHARGE RELAY SYSTEM.
FT   ACT_SITE     197    197       CHARGE RELAY SYSTEM.
FT   SITE         191    191       REQUIRED FOR SPECIFICITY.
FT   STRAND        22     22
FT   STRAND        25     26
FT   TURN          29     30
FT   TURN          33     34
FT   STRAND        35     39
FT   STRAND        43     51
FT   TURN          52     53
FT   STRAND        54     57
FT   HELIX         59     61
FT   STRAND        67     70
FT   TURN          75     76
FT   STRAND        83     92
FT   TURN          94     95
FT   STRAND        97     97
FT   TURN          98    101
FT   STRAND       102    102
FT   TURN         103    103
FT   STRAND       106    110
FT   STRAND       124    124
FT   TURN         132    133
FT   STRAND       135    140
FT   STRAND       156    162
FT   HELIX        165    171
FT   TURN         173    174
```

```
FT   PROPEP       17     25       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN        26    248       TRYPSIN II-P29.
FT   ACT_SITE     65     65       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    109    109       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    202    202       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     32    162       BY SIMILARITY.
FT   DISULFID     50     66       BY SIMILARITY.
FT   DISULFID    134    235       BY SIMILARITY.
FT   DISULFID    141    208       BY SIMILARITY.
FT   DISULFID    173    187       BY SIMILARITY.
FT   DISULFID    198    222       BY SIMILARITY.
FT   SITE        196    196       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   248 AA;  26622 MW;  E5E16B07622B588E CRC64;


  Query Match             61.6%;  Score 85;  DB 1;  Length 248;
  Best Local Similarity   52.0%;  Pred. No. 2.8e-06;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      44 NSGYHFCGGSLINSQWVLSAAHCYK 68


RESULT   6
GRAK_RAT
ID   GRAK_RAT       STANDARD;      PRT;   258 AA.
AC   P49864;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   GRANZYME K PRECURSOR (EC 3.4.21.-) (NK-TRYPTASE-2) (NK-TRYP-2).
GN   GZMK.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 26-58.
RC   STRAIN=FISCHER 344; TISSUE=Lymphocytes;
RX   MEDLINE=94179809; PubMed=8133042;
RA   Sayers T.J., Wiltrout T.A., Smyth M.J., Ottaway K.S., Pilaro A.M.,
RA   Sowder R., Henderson L.E., Sprenger H., Lloyd A.R.;
RT   "Purification and cloning of a novel serine protease, RNK-Tryp-2,
RT   from the granules of a rat NK cell leukemia.";
RL   J. Immunol. 152:2289-2297(1994).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES.
CC   -!- TISSUE SPECIFICITY: SPEEN, LUNGS AND LIVER NONPARENCHYMAL CELLS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. STRONGEST TO OTHER GRANZYMES AND TO MAST CELL
CC       PROTEASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L19694; AAA42057.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.146; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Zymogen; Signal.
FT   SIGNAL        1     23       POTENTIAL.
FT   PROPEP       24     25       ACTIVATION PEPTIDE.
FT   CHAIN        26    258       GRANZYME K.
FT   ACT_SITE     66     66       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    110    110       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    208    208       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     51     67       BY SIMILARITY.
FT   DISULFID    143    214       BY SIMILARITY.
FT   DISULFID    175    193       BY SIMILARITY.
FT   DISULFID    204    228       BY SIMILARITY.
SQ   SEQUENCE   258 AA;  28465 MW;  988AD71DB08AFBB4 CRC64;


  Query Match             61.6%;  Score 85;  DB 1;  Length 258;
  Best Local Similarity   56.5%;  Pred. No. 2.9e-06;
  Matches   13;  Conservative    6;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCF 24
           :|:| ||| ||| ::|:||| |:
Db      46 RGKHICGGVLIHPQWVLTAAHCY 68


RESULT   7
TRYT_MERUN
ID   TRYT_MERUN     STANDARD;      PRT;   270 AA.
AC   P50342;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MAST CELL TRYPTASE PRECURSOR (EC 3.4.21.59).
OS   Meriones unguiculatus (Mongolian jird).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC   Meriones.
OX   NCBI_TaxID=10047;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGS/SEA; TISSUE=Intestine;
RX   MEDLINE=95366971; PubMed=7639711;
RA   Murakumo Y., Ide H., Itoh H., Tomita M., Kobayashi T.,
RA   Maruyama H., Horii Y., Nawa Y.;
RT   "Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil,
RT   Meriones unguiculatus, and its preferential expression in the
RT   intestinal mucosa.";
RL   Biochem. J. 309:921-926(1995).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-|-, LYS-|-, BUT
CC       WITH MORE RESTRICTED SPECIFICITY THAN TRYPSIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. STRONGEST TO OTHER TRYPTASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D31789; BAA06598.1; -.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.143; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    270       MAST CELL TRYPTASE.
FT   ACT_SITE     69     69       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    116    116       CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```
RX   MEDLINE=90332035; PubMed=2377228;
RA   Jenne D.E., Tschopp J., Luedemann J., Utecht B., Gross W.L.;
RT   "Wegener's autoantigen decoded.";
RL   Nature 346:520-520(1990).
RN   [10]
RP   IDENTITY OF WEGENER'S AUTOANTIGEN WITH PROTEINASE 3.
RX   MEDLINE=91055123; PubMed=2242436;
RA   Gupta S.K., Niles J.L., McCluskey R.T., Arnaout M.A.;
RT   "Identity of Wegener's autoantigen (p29) with proteinase 3 and
RT   myeloblastin.";
RL   Blood 76:2162-2162(1990).
RN   [11]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS).
RX   MEDLINE=96346178; PubMed=8757293;
RA   Fujinaga M., Charnaia M.M., Halenbeck R., Koths K., James M.N.G.;
RT   "The crystal structure of PR3, a neutrophil serine proteinase antigen
RT   of Wegener's granulomatosis antibodies.";
RL   J. Mol. Biol. 261:267-278(1996).
CC   -!- FUNCTION: POLYMORPHONUCLEAR LEUKOCYTE SERINE PROTEASE THAT
CC       DEGRADES ELASTIN, FIBRONECTIN, LAMININ, VITRONECTIN, AND COLLAGEN
CC       TYPES I, III, AND IV (IN VITRO) AND CAUSES EMPHYSEMA WHEN
CC       ADMINISTERED BY TRACHEAL INSUFFLATION TO HAMSTERS.
CC   -!- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS, INCLUDING ELASTIN, BY
CC       PREFERENTIAL CLEAVAGE: ALA-|-XAA > VAL-|-XAA.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. ELASTASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X56132; CAA39598.1; -.
DR   EMBL; AC004799; AAC18958.1; -.
DR   EMBL; M75154; AAA59558.1; -.
DR   EMBL; M96839; AAB59493.1; -.
DR   EMBL; M96838; AAB59493.1; JOINED.
DR   EMBL; M96837; AAB59493.1; JOINED.
DR   EMBL; X55668; CAA39203.1; -.
DR   EMBL; M29142; AAA36342.1; -.
DR   EMBL; M96628; AAB59364.1; -.
DR   PIR; A43983; PRHU3.
DR   PDB; 1FUJ; 11-JUL-96.
DR   MEROPS; S01.134; -.
DR   MIM; 177020; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KW   3D-structure.
FT   SIGNAL       1     25
FT   PROPEP       26    27
FT   CHAIN        28    248       MYELOBLASTIN.
FT   PROPEP       249   256
FT   ACT_SITE     71    71        CHARGE RELAY SYSTEM.
FT   ACT_SITE     118   118       CHARGE RELAY SYSTEM.
FT   ACT_SITE     203   203       CHARGE RELAY SYSTEM.
FT   CARBOHYD     129   129       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD     174   174       N-LINKED (GLCNAC. . .).
FT   DISULFID     56    72
FT   DISULFID     152   209
FT   DISULFID     182   188
FT   DISULFID     199   224
FT   CONFLICT     2     2         A -> R (IN REF. 3).
FT   CONFLICT     46    46        Q -> E (IN REF. 7 AND 8).
FT   CONFLICT     64    64        S -> D (IN REF. 6).
FT   CONFLICT     70    70        A -> P (IN REF. 1).
FT   CONFLICT     119   119       V -> I (IN REF. 1 AND 5).
FT   CONFLICT     135   136       AT -> TS (IN REF. 1 AND 5).
FT   CONFLICT     255   255       MISSING (IN REF. 3).
SQ   SEQUENCE   256 AA;  27807 MW;  CBECA36D8C4B2A40 CRC64;


  Query Match             67.4%;  Score 93;  DB 1;  Length 256;
  Best Local Similarity   64.0%;  Pred. No. 1.6e-07;
  Matches   16;  Conservative    2;  Mismatches    7;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | ||||| |||  ||:||| | :
Db      50 NPGSHFCGGTLIHPSFVLTAAHCLR 74


RESULT   4
TRYP_PIG
ID   TRYP_PIG       STANDARD;      PRT:   231 AA.
AC   P00761;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   TRYPSIN PRECURSOR (EC 3.4.21.4).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE OF 1-10.
RA   Charles M., Rovery M., Guidoni A.A., Desnuelle P.;
RT   "On trypsinogen and trypsin of pig.";
RL   Biochim. Biophys. Acta 69:115-129(1963).
RN   [2]
RP   SEQUENCE OF 9-231.
RX   MEDLINE=73258692; PubMed=4738933;
RA   Hermodson M.A., Ericsson L.H., Neurath H., Walsh K.A.;
RT   "Determination of the amino acid sequence of porcine trypsin by
RT   sequenator aalysis.";
RL   Biochemistry 12:3146-3153(1973).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX   MEDLINE=93187998; PubMed=8445634;
RA   Huang Q., Liu S., Tang Y.;
RT   "Refined 1.6-A resolution crystal structure of the complex formed
RT   between porcine beta-trypsin and MCTI-A, a trypsin inhibitor of the
RT   squash family. Detailed comparison with bovine beta-trypsin and its
RT   complex.";
RL   J. Mol. Biol. 229:1022-1030(1993).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS).
RX   MEDLINE=92201369; PubMed=1551419;
RA   Huang Q., Liu S., Tang Y., Zeng F., Qian R.;
RT   "Amino acid sequencing of a trypsin inhibitor by refined 1.6 A X-ray
RT   crystal structure of its complex with porcine beta-trypsin.";
RL   FEBS Lett. 297:143-146(1992).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX   MEDLINE=95035057; PubMed=7947985;
RA   Huang Q., Wang Z., Li Y., Liu S., Tang Y.;
RT   "Refined 1.8-A resolution crystal structure of the porcine epsilon-
RT   trypsin.";
RL   Biochim. Biophys. Acta 1209:77-82(1994).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) IN COMPLEX WITH LDTI.
RX   MEDLINE=97390427; PubMed=9242660;
RA   Stubbs M.T., Morenweiser R., Sturzebecher J., Bauer M., Bode W.,
RA   Huber R., Piechottka G.P., Matschiner G., Sommerhoff C.P., Fritz H.,
RA   Auerswald E.A.;
RT   "The three-dimensional structure of recombinant leech-derived
RT   tryptase inhibitor in complex with trypsin. Implications for the
RT   structure of human mast cell tryptase and its inhibition.";
```

```
RT   antibacterial and monocyte-specific chemotactic glycoprotein
RT   structurally similar to neutrophil elastase.";
RL   FEBS Lett. 272:200-204(1990).
RN   [6]
RP   SEQUENCE OF 27-248.
RC   TISSUE=Neutrophils;
RX   MEDLINE=91224149; PubMed=2026172;
RA   Flodgaard H., Oestergaard E., Bayne S., Svendsen A., Thomsen J.,
RA   Engels M., Wollmer A.;
RT   "Covalent structure of two novel neutrophile leucocyte-derived
RT   proteins of porcine and human origin. Neutrophile elastase homologues
RT   with strong monocyte and fibroblast chemotactic activities.";
RL   Eur. J. Biochem. 197:535-547(1991).
RN   [7]
RP   SEQUENCE OF 27-47.
RX   MEDLINE=91378304; PubMed=1897955;
RA   Green B.G., Weston H., Ashe B.M., Doherty J., Finke P., Hagmann W.,
RA   Lark M., Mao J., Maycock A., Moore V., Mumford R., Shah S.,
RA   Walakovits L., Knight W.B.;
RT   "PMN elastases: a comparison of the specificity of human isozymes and
RT   the enzyme from other species toward substrates and inhibitors.";
RL   Arch. Biochem. Biophys. 286:284-292(1991).
RN   [8]
RP   SEQUENCE OF 27-46.
RX   MEDLINE=89315847; PubMed=2501794;
RA   Gabay J.E., Scott R.W., Campanelli D., Griffith J., Wilde C.,
RA   Marra M.N., Seeger M., Nathan C.F.;
RT   "Antibiotic proteins of human polymorphonuclear leukocytes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:5610-5614(1989).
RN   [9]
RP   SEQUENCE OF 27-46 AND 194-217.
RX   MEDLINE=90130450; PubMed=2404977;
RA   Wilde C.G., Snable J.L., Griffith J.E., Scott R.W.;
RT   "Characterization of two azurphil granule proteases with active-site
RT   homology to neutrophil elastase.";
RL   J. Biol. Chem. 265:2038-2041(1990).
RN   [10]
RP   SEQUENCE OF 27-67.
RC   TISSUE=Neutrophils;
RX   MEDLINE=90157837; PubMed=2406527;
RA   Pereira H.A., Spitznagel J.K., Pohl J., Wilson D.E., Morgan J.,
RA   Palings I., Larrick J.W.;
RT   "CAP 37, a 37 kD human neutrophil granule cationic protein shares
RT   homology with inflammatory proteinases.";
RL   Life Sci. 46:189-196(1990).
RN   [11]
RP   SEQUENCE OF 27-68.
RC   TISSUE=Neutrophils;
RX   MEDLINE=90237224; PubMed=2332502;
RA   Pereira H.A., Shafer W.M., Pohl J., Martin L.E., Spitznagel J.K.;
RT   "CAP37, a human neutrophil-derived chemotactic factor with monocyte
RT   specific activity.";
RL   J. Clin. Invest. 85:1468-1476(1990).
RN   [12]
RP   REVIEW.
RX   MEDLINE=92095171; PubMed=1755363;
RA   Morgan J.G., Pereira H.A., Sukiennicki T., Spitznagel J.K.,
RA   Larrick J.W.;
RT   "Human neutrophil granule cationic protein CAP37 is a specific
RT   macrophage chemotaxin that shares homology with inflammatory
RT   proteinases.";
RL   Adv. Exp. Med. Biol. 305:89-96(1991).
RN   [13]
RP   SYNTHESIS OF 46-70.
RX   MEDLINE=93281653; PubMed=8506327;
RA   Pereira H.A., Erdem I., Pohl J., Spitznagel J.K.;
RT   "Synthetic bactericidal peptide based on CAP37: a 37-kDa human
RT   neutrophil granule-associated cationic antimicrobial protein
RT   chemotactic for monocytes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:4733-4737(1993).
RN   [14]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=97249288; PubMed=9095193;
RA   Iversen L.F., Kastrup J.S., Bjoern S.E., Rasmussen P.B., Wiberg F.C.,
RA   Flodgaard H.J., Larsen I.K.;
RT   "Structure of HBP, a multifunctional protein with a serine proteinase
RT   fold.";
RL   Nat. Struct. Biol. 4:265-268(1997).
RN   [15]
RP   X-RAY CRYSTALLOGRAPHY (1.12 ANGSTROMS).
RX   MEDLINE=98437573; PubMed=9761855;
RA   Karlsen S., Iversen L.F., Larsen I.K., Flodgaard H.J., Kastrup J.S.;
RT   "Atomic resolution structure of human HBP/CAP37/azurocidin.";
RL   Acta Crystallogr. D 54:598-609(1998).
CC   -!- FUNCTION: THIS IS A NEUTROPHIL GRANULE-DERIVED ANTIBACTERIAL AND
CC       MONOCYTE- AND FIBROBLAST-SPECIFIC CHEMOTACTIC GLYCOPROTEIN. BINDS
CC       HEPARIN. THE CYTOTOXIC ACTION IS LIMITED TO MANY SPECIES OF GRAM-
CC       NEGATIVE BACTERIA; THIS SPECIFICITY MAY BE EXPLAINED BY A STRONG
CC       AFFINITY OF THE VERY BASIC N-TERMINAL HALF FOR THE NEGATIVELY
CC       CHARGED LIPOPOLYSACCHARIDES THAT ARE UNIQUE TO THE GRAM-NEGATIVE
CC       BACTERIAL OUTER ENVELOPE. IT MAY PLAY A ROLE IN MEDIATING
CC       RECRUITMENT OF MONOCYTES IN THE SECOND WAVE OF INFLAMMATION.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC GRANULES OF NEUTROPHILS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. ELASTASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96326; AAB59353.1; -.
DR   EMBL; X58794; CAA41601.1; -.
DR   EMBL; AC004799; AAC18957.1; -.
DR   PIR; A46455; TRHUAZ.
DR   PIR; A46268; A46268.
DR   PDB; 1AE5; 11-MAR-98.
DR   PDB; 1A7S; 23-MAR-99.
DR   MEROPS; S01.971; -.
DR   MIM; 162815; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW   Serine protease homolog; Glycoprotein; Chemotaxis; Antibiotic;
KW   Heparin-binding; Signal; 3D-structure.
FT   SIGNAL       1     26
FT   CHAIN       27    248       AZUROCIDIN.
FT   PROPEP     249    251
FT   DOMAIN      46     70       POSSESS ANTIBIOTIC ACTIVITY.
FT   DOMAIN      52     84       HYDROPHOBIC.
FT   CARBOHYD   126    126       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   CARBOHYD   140    140       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   171    171       N-LINKED (GLCNAC. . .) (PARTIAL).
FT   DISULFID    52     68
FT   DISULFID   149    207
FT   DISULFID   180    186
FT   DISULFID   197    222
FT   VARIANT    248    248       MISSING (IN 50% OF THE MOLECULES).
FT                               /FTId=VAR_006496.
FT   MUTAGEN     52     52       C->S: LOSS OF ANTIBIOTIC ACTIVITY.
FT   MUTAGEN     68     68       C->S: LOSS OF ANTIBIOTIC ACTIVITY.
FT   CONFLICT    36     36       R -> H (IN REF. 9).
FT   CONFLICT   130    130       S -> N (IN REF. 6).
SQ   SEQUENCE   251 AA;  26885 MW;  22F80D9EBE87DE60 CRC64;


 Query Match           100.0%;  Score 138;  DB 1;  Length 251;
```

```
RESULT 13
S55067
trypsin (EC 3.4.21.4) I precursor, pancreatic - chicken
N;Alternate names: trypsinogen I
C;Species: Gallus gallus (chicken)
C;Date: 28-Oct-1996 #sequence_revision 07-Feb-1997 #text_change 21-Jul-2000
C;Accession: S55067; S72345; S55065; S72346; S71155
R;Wang, K.; Gan, L.; Lee, I.; Hood, L.
Biochem. J. 307, 471-479, 1995
A;Title: Isolation and characterization of the chicken trypsinogen gene family.
A;Reference number: S55065; MUID:95251611
A;Accession: S55067
A;Molecule type: mRNA
A;Residues: 1-248 <WAN1>
A;Cross-references: EMBL:U15156; NID:g603904
A;Experimental source: clone 1-P38
A;Accession: S72345
A;Molecule type: DNA
A;Residues: 1-248 <WAN2>
A;Experimental source: clone 1-P38
A;Accession: S55065
A;Molecule type: mRNA
A;Residues: 1-9,'V',11-12,'T',14-102,'A',104-214,'I',216-248 <WAN3>
A;Cross-references: EMBL:U15155; NID:g603902; PIDN:AAA79912.1; PID:g603903
A;Experimental source: clone P1
A;Accession: S72346
A;Molecule type: DNA
A;Residues: 1-9,'V',11-12,'T',14-102,'A',104-214,'I',216-248 <WAN4>
A;Cross-references: GB:U15155; NID:g603902; PIDN:AAA79912.1; PID:g603903
A;Experimental source: clone P1
R;Wang, K.
submitted to the EMBL Data Library, September 1994
A;Reference number: S71155
A;Accession: S71155
A;Molecule type: mRNA
A;Residues: 1-102,'A',104-248 <WAN5>
A;Cross-references: EMBL:U15156; NID:g603904; PIDN:AAA79913.1; PID:g603905
A;Experimental source: clone 1-P38
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-25/Domain: activation peptide #status predicted <APT>
F;26-248/Product: trypsin I #status predicted <MAT>
F;26-241/Domain: trypsin homology <TRY>
F;65,109,202/Active site: His, Asp, Ser #status predicted


  Query Match             60.1%;  Score 83;  DB 2;  Length 248;
  Best Local Similarity   52.0%;  Pred. No. 4.8e-05;
  Matches   13;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:|| :::|::|| |::
Db      44 NSGYHFCGGSLISSQWVLSAAHCYK 68


RESULT 14
A47136
macrophage-stimulating protein 1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-May-1994 #sequence_revision 14-Nov-1997 #text_change 18-Jun-1999
C;Accession: A40331; B40331; A47136; A61395
R;Han, S.; Stuart, L.A.; Degen, S.J.F.
Biochemistry 30, 9768-9780, 1991
A;Title: Characterization of the DNF15S2 locus on human chromosome 3: identification of
A;Reference number: A40331; MUID:92002016
A;Accession: A40331
A;Molecule type: DNA
A;Residues: 1-711 <HA1>
A;Cross-references: GB:M74179
A;Accession: B40331
A;Molecule type: mRNA
A;Residues: 1-711 <HA2>
A;Cross-references: GB:M74178; NID:g183976; PIDN:AAA50165.1; PID:g183977
R;Yoshimura, T.; Yuhki, N.; Wang, M.H.; Skeel, A.; Leonard, E.J.
J. Biol. Chem. 268, 15461-15468, 1993
A;Title: Cloning, sequencing, and expression of human macrophage stimulating protein
A;Reference number: A47136; MUID:93340141
A;Accession: A47136
A;Molecule type: mRNA
A;Residues: 1-12,'C',14-622,'F',624-711 <YOS>
A;Cross-references: GB:L11924; NID:g398037; PIDN:AAA59872.1; PID:g398038
A;Note: authors translated the codon TTT for residue 623 as Leu; parts of this sequen
R;Skeel, A.; Yoshimura, T.; Showalter, S.D.; Tanaka, S.; Appella, E.; Leonard, E.J.
J. Exp. Med. 173, 1227-1234, 1991
A;Title: Macrophage stimulating protein: purification, partial amino acid sequence, a
A;Reference number: A61395; MUID:91217635
A;Accession: A61395
A;Molecule type: protein
A;Residues: 230-247;288-291,'E',293-295,'X',297-301,'X',303,'E',305,'EX',308-310;326-
A;Experimental source: plasma
C;Genetics:
A;Gene: GDB:MST1; D3P15S2; DNF15S2; HGFL
A;Cross-references: GDB:128833; OMIM:142408
A;Map position: 3p21-3p21.3
C;Complex: disulfide-bonded heterodimer of chains derived from the same precursor
C;Superfamily: hepatocyte growth factor; kringle homology; trypsin homology
C;Keywords: duplication; glycoprotein; growth factor; kringle; plasma
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-483,484-711/Product: macrophage-stimulating protein 1 #status predicted <MAT>
F;19-483/Domain: alpha chain #status predicted <ACH>
F;110-186/Domain: kringle homology <KR1>
F;191-268/Domain: kringle homology <KR2>
F;283-361/Domain: kringle homology <KR3>
F;370-448/Domain: kringle homology <KR4>
F;484-711/Domain: beta chain #status predicted <BCH>
F;484-704/Domain: trypsin homology <TRY>
F;56-78,60-66,110-186,131-169,157-181,191-268,212-251,240-263,283-361,304-343,332-355
F;72,296,615/Binding site: carbohydrate (Asn) (covalent) #status predicted


  Query Match             60.1%;  Score 83;  DB 1;  Length 711;
  Best Local Similarity   52.2%;  Pred. No. 0.00012;
  Matches   12;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCF 24
           ||:|||||:|:  ::::||  ||
Db     502 QGQHFCGGSLVKEQWILTARQCF 524


RESULT 15
A61545
plasmin (EC 3.4.21.7) precursor - horse (fragments)
N;Alternate names: plasminogen
N;Contains: miniplasminogen
C;Species: Equus caballus (domestic horse)
C;Date: 28-Oct-1994 #sequence_revision 01-Nov-1996 #text_change 18-Jul-1997
C;Accession: A61545; S17527
R;Schaller, J.; Rickli, E.E.
Enzyme 40, 63-69, 1988
A;Title: Structural aspects of the plasminogen of various species.
A;Reference number: A61545; MUID:89005015
A;Accession: A61545
A;Molecule type: protein
A;Residues: 1-33;34-117 <SCH>
R;Schaller, J.; Straub, C.; Kaempfer, U.; Rickli, E.E.
Protein Seq. Data Anal. 4, 69-74, 1991
A;Title: Complete amino acid sequence of equine miniplasminogen.
A;Reference number: S17527; MUID:92052077
A;Accession: S17527
A;Molecule type: protein
A;Residues: 118-455 <SC2>
C;Superfamily: plasmin; kringle homology; plasminogen-related protein precursor homol
C;Keywords: fibrinolysis; glycoprotein; hydrolase; kringle; plasma; serine proteinase
F;1-33,34-117,118-455/Product: plasminogen (fragments) #status experimental <PRO>
```

F;9-224/Domain: trypsin homology <TRY>
F;15-145,33-49,117-218,124-191,156-170,181-205/Disulfide bonds: #status predicted
F;48,92,185/Active site: His, Asp, Ser #status predicted
F;60,62,65,70/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Query Match 61.6%; Score 85; DB 1; Length 231;
Best Local Similarity 52.0%; Pred. No. 2.3e-05;
Matches 13; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
Db 27 NSGSHFCGGSLINSQWVVSAAHCYK 51

RESULT 5
S55066
trypsin (EC 3.4.21.4) II precursor, pancreatic (clone 2-P29) - chicken
N;Alternate names: trypsinogen II
C;Species: Gallus gallus (chicken)
C;Date: 23-Aug-1995 #sequence_revision 19-Oct-1995 #text_change 22-Jun-1999
C;Accession: S55066; S72347
R;Wang, K.; Gan, L.; Lee, I.; Hood, L.
Biochem. J. 307, 471-479, 1995
A;Title: Isolation and characterization of the chicken trypsinogen gene family.
A;Reference number: S55065; MUID:95251611
A;Accession: S55066
A;Molecule type: mRNA
A;Residues: 1-248 <WAN1>
A;Cross-references: EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g603907
A;Experimental source: clone 2-P29
A;Accession: S72347
A;Molecule type: DNA
A;Residues: 1-248 <WAN2>
A;Cross-references: EMBL:U15157; NID:g603906; PIDN:AAA79914.1; PID:g603907
A;Experimental source: clone 2-P29
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-25/Domain: activation peptide #status predicted <APT>
F;26-248/Product: trypsin II #status predicted <MAT>
F;26-241/Domain: trypsin homology <TRY>
F;65,109,202/Active site: His, Asp, Ser #status predicted

Query Match 61.6%; Score 85; DB 2; Length 248;
Best Local Similarity 52.0%; Pred. No. 2.4e-05;
Matches 13; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
Db 44 NSGYHFCGGSLINSQWVLSAAHCYK 68

RESULT 6
I56220
tryptase 2 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 22-Jun-1999
C;Accession: I56220
R;Sayers, T.J.; Wiltrout, T.A.; Smyth, M.J.; Ottaway, K.S.; Pilaro, A.M.; Sowder, R.; He
J. Immunol. 152, 2289-2297, 1994
A;Title: Purification and cloning of a novel serine protease, RNK-Tryp-2, from the granu
A;Reference number: I56220; MUID:94179809
A;Accession: I56220
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-258 <RES>
A;Cross-references: GB:L19694; NID:g487878; PIDN:AAA42057.1; PID:g487879
C;Superfamily: trypsin; trypsin homology
F;26-248/Domain: trypsin homology <TRY>

Query Match 61.6%; Score 85; DB 2; Length 258;
Best Local Similarity 56.5%; Pred. No. 2.5e-05;
Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Qy 2 QGRHFCGGALIHARFVMTAASCF 24
Db 46 RGKHICGGVLIHPQWVLTAAHCY 68

RESULT 7
S56160
mast cell tryptase precursor - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C;Accession: S56160
R;Murakumo, Y.; Ide, H.; Itoh, H.; Tomita, M.; Kobayashi, T.; Maruyama, H.; Horii, Y.
Biochem. J. 309, 921-926, 1995
A;Title: Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil, Merione
A;Reference number: S56160; MUID:95366971
A;Accession: S56160
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-270 <MUR>
A;Cross-references: EMBL:D31789; NID:g517122; PIDN:BAA06598.1; PID:g517123
C;Superfamily: trypsin; trypsin homology
F;26-262/Domain: trypsin homology <TRY>

Query Match 61.6%; Score 85; DB 2; Length 270;
Best Local Similarity 70.0%; Pred. No. 2.6e-05;
Matches 14; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RHFCGGALIHARFVMTAASC 23
Db 51 RHFCGGSLIHPQWVLTAAHC 70

RESULT 8
TRBOTR
trypsin (EC 3.4.21.4) precursor - bovine
N;Contains: trypsinogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change 18-Jul-1997
C;Accession: A90164; A00946; S08774
R;Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
Biochem. Biophys. Res. Commun. 24, 346-352, 1966
A;Title: Covalent structure of bovine trypsinogen. The position of the remaining amid
A;Reference number: A90164; MUID:67168848
A;Accession: A90164
A;Molecule type: protein
A;Residues: 1-57,'Q',59-67,'Q',69-150,'N',152-176,'N',178-229 <MIK>
R;Hartley, B.S.
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970
A;Reference number: A93755
A;Contents: annotation; revisions
R;Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975
A;Title: Amino acid sequence of dogfish trypsin.
A;Reference number: A00950; MUID:75146445
A;Contents: annotation; revisions
A;Note: the sequence agrees with that shown
R;Bode, W.; Schwager, P.
J. Mol. Biol. 98, 693-717, 1975
A;Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolut
A;Reference number: A92954; MUID:76072097
A;Contents: annotation; X-ray crystallography; binding sites for calcium, substrate,
C;Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.
C;Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a te
s pseudotrypsin. A cleavage may also occur after Arg-105.
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-229/Product: trypsinogen #status experimental <ZYM>
F;1-6/Domain: activation peptide #status experimental <APT>

R;Green, B.G.; Weston, H.; Ashe, B.M.; Doherty, J.; Finke, P.; Hagmann, W.; Lark, M.; Ma
Arch. Biochem. Biophys. 286, 284-292, 1991
A;Title: PMN elastases: a comparison of the specificity of human isozymes and the enzyme
A;Reference number: S14736; MUID:91378304
A;Accession: S14738
A;Molecule type: protein
A;Residues: 27-47 <GRE>
R;Gabay, J.E.; Scott, R.W.; Campanelli, D.; Griffith, J.; Wilde, C.; Marra, M.N.; Seeger
Proc. Natl. Acad. Sci. U.S.A. 86, 5610-5614, 1989
A;Title: Antibiotic proteins of human polymorphonuclear leukocytes.
A;Reference number: A33913; MUID:89315847
A;Accession: B33913
A;Molecule type: protein
A;Residues: 27-46 <GAB>
R;Pereira, H.A.; Shafer, W.M.; Pohl, J.; Martin, L.E.; Spitznagel, J.K.
J. Clin. Invest. 85, 1468-1476, 1990
A;Title: CAP37, a human neutrophil-derived chemotactic factor with monocyte specific act
A;Reference number: A60708; MUID:90237224
A;Accession: A60708
A;Molecule type: protein
A;Residues: 27-68 <PER>
R;Wasiluk, K.R.; Skubitz, K.M.; Gray, B.H.
Infect. Immun. 59, 4193-4200, 1991
A;Title: Comparison of granule proteins from human polymorphonuclear leukocytes which ar
A;Reference number: A43600; MUID:92040097
A;Accession: B43600
A;Molecule type: protein
A;Residues: 27-48 <WAS>
R;Miyasaki, K.T.; Bodeau, A.L.
Infect. Immun. 60, 4973-4975, 1992
A;Title: Human neutrophil azurocidin synergizes with leukocyte elastase and cathepsin G
A;Reference number: A49211; MUID:93014226
A;Accession: A49211
A;Molecule type: protein
A;Residues: 27-46 <MIY>
A;Note: sequence extracted from NCBI backbone (NCBIP:116551)
R;Shellard, J.E.; Leitch, H.A.; Logan, P.M.; McMaster, W.R.; Levy, J.G.
Exp. Hematol. 19, 136-142, 1991
A;Title: Purification of an in vitro inhibitor of normal myelopoiesis using a monoclonal
A;Reference number: A61502; MUID:91122218
A;Accession: A61502
A;Molecule type: protein
A;Residues: 27-48 <SHE>
R;Wilde, C.G.; Snable, J.L.; Griffith, J.E.; Scott, R.W.
J. Biol. Chem. 265, 2038-2041, 1990
A;Title: Characterization of two azurophil granule proteases with active-site homology t
A;Reference number: A43981; MUID:90130450
A;Accession: B43981
A;Molecule type: protein
A;Residues: 27-35,'H',37-46;194-217 <WIL>
R;Pereira, H.A.; Spitznagel, J.K.; Pohl, J.; Wilson, D.E.; Morgan, J.; Palings, I.; Larr
Life Sci. 46, 189-196, 1990
A;Title: CAP 37, a 37 kD human neutrophil granule cationic protein shares homology with
A;Reference number: PH0081; MUID:90157837
A;Accession: PH0081
A;Molecule type: protein
A;Residues: 27-67 <PE2>
A;Experimental source: polymorphonuclear leukocyte
C;Comment: This protein is homologous to serine proteinases but lacks proteolytic activi
icrobial activity against a number of gram-negative bacteria.
C;Genetics:
A;Gene: GDB:AZU1
A;Cross-references: GDB:135033; OMIM:162815
A;Map position: 19p13.3-19p13.3
A;Introns: 20/1; 72/2; 150/3; 198/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; inflammation
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-26/Domain: amino-terminal propeptide #status predicted <PRO>
F;27-248/Product: azurocidin #status experimental <MAT>
F;27-239/Domain: trypsin homology <TRY>
F;67,115,201/Region: defective catalytic triad
F;249-251/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;52-68/Disulfide bonds: #status experimental
F;126,140,171/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;149-207,180-186,197-222/Disulfide bonds: #status predicted

Query Match 100.0%; Score 138; DB 1; Length 251;
Best Local Similarity 100.0%; Pred. No. 3.1e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
||||||||||||||||||||||||||
Db 46 NQGRHFCGGALIHARFVMTAASCFQ 70

RESULT 2
TRPGAZ
azurocidin - pig
N;Alternate names: heparin-binding protein
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 18-Jul-1997
C;Accession: S15393
R;Flodgaard, H.; Ostergaard, E.; Bayne, S.; Svendsen, A.; Thomsen, J.; Engels, M.; Wo
Eur. J. Biochem. 197, 535-547, 1991
A;Title: Covalent structure of two novel neutrophile leucocyte-derived proteins of po
.
A;Reference number: S15393; MUID:91224149
A;Accession: S15393
A;Molecule type: protein
A;Residues: 1-219 <FLO>
R;Sorensen, H.H.; Thomsen, J.; Bayne, S.; Hojrup, P.; Roepstorff, P.
Biomed. Environ. Mass Spectrom. 19, 713-720, 1990
A;Title: Strategies for determination of disulphide bridges in proteins using plasma
A;Reference number: A37448; MUID:91167744
A;Contents: annotation; disulfide bonds
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein
F;1-211/Domain: trypsin homology <TRY>
F;41,88,173/Region: defective catalytic triad
F;26-42,122-179,152-158/Disulfide bonds: #status experimental
F;113,144/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 74.6%; Score 103; DB 1; Length 219;
Best Local Similarity 75.0%; Pred. No. 4.6e-08;
Matches 18; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QGRHFCGGALIHARFVMTAASCFQ 25
||| || |||:| |||:|||||:
Db 21 QGRPFCAGALVHPRFVLTAASCFR 44

RESULT 3
PRHU3
proteinase 3 (EC 3.4.21.-) precursor [validated] - human
N;Alternate names: AGP7; C-ANCA antigen; neutrophil proteinase 4; p29; Wegener's gran
N;Contains: myelobiastin
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 08-Dec-2000
C;Accession: A45080; B46268; A43983; JH0331; A33751; S11091; A61176; A60148; A43982;
R;Sturrock, A.B.; Franklin, K.F.; Rao, G.; Marshall, B.C.; Rebentisch, M.B.; Lemons,
J. Biol. Chem. 267, 21193-21199, 1992
A;Title: Structure, chromosomal assignment, and expression of the gene for proteinase
A;Reference number: A45080; MUID:93016043
A;Accession: A45080
A;Molecule type: DNA
A;Residues: 1-254,'P' <STU>
A;Cross-references: GB:M97911
A;Note: sequence extracted from NCBI backbone (NCBIP:116205)
R;Zimmer, M.; Medcalf, R.L.; Fink, T.M.; Mattmann, C.; Lichter, P.; Jenne, D.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 8215-8219, 1992
A;Title: Three human elastase-like genes coordinately expressed in the myelomonocyte
A;Reference number: A46268; MUID:92390417
A;Accession: B46268

```
  Best Local Similarity   100.0%;  Pred. No. 4.3e-09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 nqgrhfcggaliharfvmtaascfq 44


RESULT  15
AAY21551
ID   AAY21551 standard; Protein; 225 AA.
XX
AC   AAY21551;
XX
DT   03-AUG-1999  (first entry)
XX
DE   Human heparin-binding protein (HBP) signal, pro and mature sequence.
XX
KW   Heparin-binding protein; HBP; apoptosis; Islets of Langerhans; beta cell;
KW   endothelial cell; nerve cell; azurophil granule; chemoattractant; cancer;
KW   polymorphonuclear leukocyte; monocyte; neurodegenerative disorder; human;
KW   neuromuscular disorder; human immunodeficiency virus; ischemic stroke;
KW   proteoglycan; mitochondrial matrix targeting protein; viral infection;
KW   apoptosis inhibitor; immune-mediated glomerulonephritis.
XX
OS   Homo sapiens.
XX
PN   WO9926647-A1.
XX
PD   03-JUN-1999.
XX
PF   20-NOV-1998;   98WO-DK00510.
XX
PR   03-DEC-1997;   97DK-0001394.
PR   20-NOV-1997;   97DK-0001324.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Flodgaard HJ;
XX
DR   WPI; 1999-347615/29.
DR   N-PSDB; AAX60695.
XX
PT   Modulating or decreasing apoptosis in cells useful for treating
PT   neurodegenerative disorders and neuromuscular disorders
XX
PS   Claim 4, 5; Page 12; 74pp; English.
XX
CC   The invention relates to a new method of modulating or decreasing
CC   apoptosis in mammalian beta cells of Islets of Langerhans, endothelial
CC   cells or nerve cells. The method comprises administering a mammalian
CC   heparin-binding protein (HBP) which in a glycosylated form has: (i) a
CC   molecular weight of 28 kDa; (ii) is produced in the azurophil granules
CC   of polymorphonuclear leukocytes; and (iii) is a chemoattractant for
CC   monocytes. The new protein is useful for preventing or treating a
CC   disorder (preferably in human patients) resulting from apoptosis of
CC   mammalian cells, including insufficient functioning of insulin production
CC   or insulin action, a neurodegenerative disorder, a neuromuscular
CC   disorder, human immunodeficiency virus and ischemic stroke. Compositions
CC   (a) comprising HBP and a proteoglycan which binds the HBP; and (b)
CC   comprising the HBP and a protein which is a mammalian mitochondrial
CC   matrix targeting protein and which binds to the HBP are useful for
CC   modulating or decreasing apoptosis in mammalian beta cells of Islets of
CC   Langerhans, endothelial cells or nerve cells. Unlike prior inhibitors,
CC   the new apoptosis inhibitor does not cause disorders associated with the
CC   inhibition of apoptosis such as cancer, autoimmune disorders such as
CC   systemic lupus erythematosus, immune-mediated glomerulonephritis, and
CC   viral infections. The present sequence represents the human HBP
CC   signal sequence, prosequence and the mature sequence.
XX
SQ   Sequence   225 AA;
```

```
  Query Match           100.0%;  Score 138;  DB 20;  Length 225;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 nqgrhfcggaliharfvmtaascfq 44


Search completed: April  5, 2002, 14:59:47
Job time: 661 sec
```

```
XX
PR   17-MAR-1988;   88DK-0001453.
XX
PA   (NORD-) NORDISK GENTOFTE A/S.
XX
PI   Flodgaard H,  Ostergaard E,  Thomsen J,  Bayne S;
XX
DR   WPI; 1989-292496/40.
XX
PT   Heparin-binding protein - having specified mol. wt. and exhibiting
PT   angiogenic properties in vivo.
XX
PS   Claim 10; page 47; 58pp; English.
XX
CC   The sequence is that of human heparin binding protein (HBP).  Amino
CC   acids X94, X98, and X131 are glycosylated and are probably Asn. Amino
CC   acids X74 and X75 are unknown.  The no. amino acids at positions 86, 106,
CC   122, and 167 is unspecified.   The HBP can be used to stimulate healing
CC   of chronic wounds, severe burns, or in tumour therapy.
CC   See also AAP90806.
XX
SQ   Sequence   197 AA;


  Query Match           100.0%;  Score 138;  DB 10;  Length 197;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 nqgrhfcggaliharfvmtaascfq 44


RESULT  11
AAR84663
ID   AAR84663 standard; protein; 221 AA.
XX
AC   AAR84663;
XX
DT   02-JUN-1996  (first entry)
XX
DE   Human heparin-binding protein.
XX
KW   Heparin-binding protein; Gram-negative sepsis; septic shock;
KW   disseminated intravascular coagulation; meningococcal meningitis;
KW   lipopolysaccharide-induced cytokine cascade disease; prevention;
KW   therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         20..53
FT                   /note= "active fragment"
FT   Peptide         26..42
FT                   /note= "active fragment"
FT   Misc-difference 100
FT                   /label= Asn_or_Asp
FT                   /note= "Asn in AAR84665 and AAR84666"
FT   Misc-difference 104
FT                   /label= Asn_or_Asp
FT                   /note= "Ser in AAR84665 and AAR84666"
FT   Misc-difference 114
FT                   /label= Asn_or_Asp
FT                   /note= "Asn in AAR84665 and AAR84666"
FT   Misc-difference 145
FT                   /label= Asn_or_Asp
FT                   /note= "Asn in AAR84665 and AAR84666"
XX
PN   WO9528949-A1.
XX
PD   02-NOV-1995.
XX
PF   17-MAR-1995;   95WO-DK00121.
XX
PR   21-DEC-1994;   94DK-0001452.
PR   21-APR-1994;   94DK-0000464.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Flodgaard HJH,  Rasmussen PB;
XX
DR   WPI; 1995-382842/49.
XX
PT   Heparin-binding protein (HBP) composition - for prevention and
PT   treatment of sepsis and other conditions.
XX
PS   Claim 3; Page 27-28; 48pp; English.
XX
CC   Heparin-binding protein (HBP) is produced by recombinant methods,
CC   where host cells containing DNA encoding mature HBP preceded by an N-
CC   terminal extension (see AAT05128/T05129) are cultured in a suitable
CC   medium under conditions permitting HBP expression.  The medium
CC   contains a sulphated polysaccharide (heparin) immobilised on an
CC   inert carrier (agarose).  To facilitate production of mature HBP,
CC   a preferred DNA encodes an N-terminally extended HBP including a
CC   protease cleavage site located between the N-terminal extension and
CC   the mature HBP coding sequence, i.e. (Asp)4Lys (enterokinase) or
CC   Ile-Glu-Gly-Arg (Factor-Xa).  HBP, or its fragments, is used in a
CC   pharmaceutical composition for the prevention or treatment of
CC   diseases associated with induction of the cytokine cascade by
CC   glycosylated lipopolysaccharide, specifically Gram-negative sepsis,
CC   septic shock, disseminated intravascular coagulation or
CC   meningococcal meningitis.
XX
SQ   Sequence   221 AA;


  Query Match           100.0%;  Score 138;  DB 16;  Length 221;
  Best Local Similarity   100.0%;  Pred. No. 4.3e-09;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      20 nqgrhfcggaliharfvmtaascfq 44


RESULT  12
AAW04875
ID   AAW04875 standard; protein; 221 AA.
XX
AC   AAW04875;
XX
DT   20-DEC-1996  (first entry)
XX
DE   Heparin binding protein.
XX
KW   Heparin binding protein; conjugate; lipid; ceramide; analogue;
KW   treatment; inflammation; viral infection; sepsis; septic shock;
KW   ischaemic reperfusion syndrome; immune system; stimulation;
KW   protein phosphatase; disseminated intravascular coagulation.
XX
OS   Homo sapiens.
XX
PN   WO9628544-A1.
XX
PD   19-SEP-1996.
XX
PF   11-MAR-1996;   96WO-DK00099.
XX
PR   09-MAR-1995;   95DK-0000240.
XX
PA   (NOVO ) NOVO-NORDISK AS.
XX
PI   Flodgaard H,  Rasmussen PB;
XX
DR   WPI; 1996-433819/43.
```

```
DE   CAP37 peptide analogue.
XX
KW   Cationic antimicrobial protein; CAP37; infection; septic shock;
KW   treatment; prevention; prophylaxis; lipopolysaccharide endotoxin;
KW   pathogen; analogue; human.
XX
OS   Homo sapiens.
XX
PN   US6107460-A.
XX
PD   22-AUG-2000.
XX
PF   01-MAR-1999;   99US-0258934.
XX
PR   01-MAR-1999;   99US-0258934.
XX
PA   (UYOK-) UNIV OKLAHOMA.
XX
PI   Pereira HA;
XX
DR   WPI; 2000-578551/54.
XX
PT   Novel peptide analogs derived from CAP37 peptides 23-42, comprising
PT   serine or threonine substitution at one of two cysteine residues at
PT   positions 26 and 42 useful for treating bacterial infections
XX
PS   Disclosure; Columns 13-14; 40pp; English.
XX
CC   Peptide derivatives based on the Cationic Antimicrobial Protein
CC   (CAP37) peptide fragment comprising amino acids 23-42 are useful
CC   for treating bacterial infections or treating or preventing septic
CC   shock in a patient. Two new analogues of a larger peptide fragment
CC   of CAP37, 20-44Ser26 and 20-44Ser26 were also active against
CC   the gram-positive bacterial pathogens, Staphylococcus aureus and
CC   Enterococcus faecalis and can bind and neutralise the toxic effects
CC   of the LPS endotoxin. The peptides 20-44Ser26 and 20-44Ser42 can
CC   also rescue mice from a lethal dose Salmonella typhimurium.
CC   The 20-44Ser26 and 20-44Ser42 analogues are easier to synthesise and
CC   purify because the internal disulphide bonding between the two
CC   cysteine residues can no longer occur. The solubility of these
CC   two new peptides is also greater than the 20-44 peptide, enabling
CC   their use at much higher concentrations than the 20-44 peptides.
CC   See GENESEQ records AAB02021-B02078.
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 21;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 nqgrhfcggaliharfvmtaascfq 25


RESULT   7
AAB09887
ID   AAB09887 standard; peptide; 25 AA.
XX
AC   AAB09887;
XX
DT   06-NOV-2000  (first entry)
XX
DE   Cationic antimicrobial protein CAP37 peptide 20-44.
XX
KW   Cationic antimicrobial protein; CAP37; antibiotic; bacterial infection;
KW   septic shock; wound healing; cancer.
XX
OS   Homo sapiens.
XX
PN   US6071879-A.
XX
PD   06-JUN-2000.
XX
PF   01-MAR-1999;   99US-0260373.
XX
PR   19-MAR-1992;   92US-0855417.
PR   07-JUN-1995;   95US-0482328.
PR   21-APR-1997;   97US-0840519.
PR   05-JUL-1989;   89US-0375739.
PR   25-JUN-1990;   90US-0543151.
PR   30-OCT-1992;   92US-0969931.
PR   29-APR-1994;   94US-0235399.
XX
PA   (OKLA ) UNIV OKLAHOMA STATE.
XX
PI   Pereira HA;
XX
DR   WPI; 2000-450766/39.
XX
PT   Treating endotoxic shock and infections caused by gram negative
PT   bacteria (e.g. Klebsiella, Pseudomonas, Escherichia coli or Salmonella)
PT   comprises administering antibiotic peptides derived from Cationic
PT   Antimicrobial Protein 37  -
XX
PS   Claim 1; Column 13-14; 29pp; English.
XX
CC   The present sequence comprises amino acid residues 20-44 of the 37kD
CC   cationic antimicrobial protein (CAP37). This protein has an antibacterial
CC   action, as the peptide is able to bind to endotoxin. Endotoxin is the
CC   major reason for bacterial pathogenesis, as the toxin can cause septic
CC   shock, the symptoms of which vary from chills and fever to circulatory
CC   failure and organ failure to death. This peptide can, therefore, be used
CC   to treat septic shock, to prevent its occurrence, and also to treat
CC   tumours and promote wound healing. The sequence is particularly useful
CC   for treating infection by gram negative bacteria such as Klebsiella,
CC   Pseudomonas, E. coli, Bacteroides and Salmonella.
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 21;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 nqgrhfcggaliharfvmtaascfq 25


RESULT   8
AAB17394
ID   AAB17394 standard; Peptide; 25 AA.
XX
AC   AAB17394;
XX
DT   31-OCT-2000  (first entry)
XX
DE   CAP37 mimetic/LPS binding peptide sequence SEQ ID NO:498.
XX
KW   Modified peptide; therapeutic agent; fusion; Fc domain; cancer;
KW   autoimmune disease; cytostatic; antiasthmatic; thrombolytic; VEGF;
KW   immunosuppressive; EPO; TPO; CTLA4; mimetic; IL-1; TNF; antagonist;
KW   MMP; inhibitor; erythropoietin; thrombopoietin; interleukin 1;
KW   cytotoxic T cell lymphocyte antigen 4; tumour necrosis factor;
KW   vascular endothelial growth factor; matrix metalloproteinase;
KW   asthma; thrombosis; pharmaceutical.
XX
OS   Synthetic.
XX
PN   WO200024782-A2.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-US25044.
```

```
XX
CC   The peptide fragment of CAP37 has chemotactic activity for monocytes.
CC   It is especially useful for treating wounds or for dental
CC   applications.  Other uses of CAP37 or its peptides include the
CC   treatment of cancer and any disease involving monocyte localisation
CC   including neoplastic diseases, parasitic diseases and rheumatoid
CC   arthritis and the treatment of endotoxaemia.  Specific CAP37
CC   peptides may be useful for treating infection due to antibacterial
CC   activity or for wound treatment because they possess antibiotic or
CC   lipopolysaccharide-binding activity.  This sequence corresponds to
CC   residues 20-44 of mature CAP37.  (See also AAR41927-41934).
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 14;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 nqgrhfcggaliharfvmtaascfq 25


RESULT   2
AAW26803
ID   AAW26803 standard; peptide; 25 AA.
XX
AC   AAW26803;
XX
DT   03-NOV-1997  (first entry)
XX
DE   Human Cationic Antimicrobial Protein CAP37 amino acids 20-44.
XX
KW   Bacteriocide; septic shock; sepsis; bacterial lipopolysacchraide;
KW   LPS; binding; systemic vascular resistance; cardiac output;
KW   tissue perfusion; white blood cell count.
XX
OS   Homo sapiens.
XX
PN   US5650392-A.
XX
PD   22-JUL-1997.
XX
PF   05-JUL-1989;   89US-0375739.
XX
PR   29-APR-1994;   94US-0235399.
PR   05-JUL-1989;   89US-0375739.
PR   25-JUN-1990;   90US-0543151.
PR   19-MAR-1992;   92US-0855417.
PR   30-OCT-1992;   92US-0969931.
PR   31-MAY-1995;   95US-0455485.
XX
PA   (OKLA ) UNIV OKLAHOMA STATE.
XX
PI   Brackett DJ,  Lerner MR,  Pereira HA;
XX
DR   WPI; 1997-384696/35.
XX
PT   Peptide(s) derived from Cationic Antimicrobial Protein CAP37 -
PT   useful for treating or preventing septic shock and for treating
PT   physiological effects induced by a bacterial lipo:polysaccharide in
PT   a mammal
XX
PS   Claim 1; Columns 9-10; 18pp; English.
XX
CC   The present sequence corresponds to amino acid residues 20-44 from
CC   a naturally occurring Cationic Antimicrobial Protein of molecular
CC   weight 37 kD. The protein, designated CAP37, can be isolated from
CC   human neutrophils. The CAP37 (20-44) peptide is useful for treating
CC   or preventing septic shock. It can also be used for treating
CC   physiological effects induced by bacterial lipopolysaccharide in a
CC   mammal, e.g. on systemic vascular resistance, cardiac output, tissue
CC   perfusion, and white blood cell count. The peptide has enhanced
CC   bacteriocidal activity compared to the native CAP37 and is active
CC   against gram-negative and gram-positive bacteria.
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 18;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 nqgrhfcggaliharfvmtaascfq 25


RESULT   3
AAW15418
ID   AAW15418 standard; peptide; 25 AA.
XX
AC   AAW15418;
XX
DT   05-JUN-1997  (first entry)
XX
DE   CAP37 fragment, 20-44.
XX
KW   Cationic antimicrobial protein; CAP37; bacterial lipopolysaccharide;
KW   LPS; septic shock; systemic vascular resistance; cardiac output;
KW   tissue perfusion; LPS-induced endotoxic response; antimicrobial;
KW   white blood cell count.
XX
OS   Synthetic.
XX
PN   US5607916-A.
XX
PD   04-MAR-1997.
XX
PF   05-JUL-1989;   89US-0375739.
XX
PR   29-APR-1994;   94US-0235399.
PR   05-JUL-1989;   89US-0375739.
PR   25-JUN-1990;   90US-0543151.
PR   19-MAR-1992;   92US-0855417.
PR   30-OCT-1992;   92US-0969931.
XX
PA   (OKLA ) UNIV OKLAHOMA STATE.
XX
PI   Brackett DJ,  Lerner MR,  Pereira HA;
XX
DR   WPI; 1997-164534/15.
XX
PT   Treating and preventing septic shock - by admin. of LPS-binding
PT   CAP37 protein fragment
XX
PS   Claim 2; Column 9-10; 19pp; English.
XX
CC   The sequences given in AAW15418-19 represent fragments.derived from
CC   cationic antimicrobial protein (CAT)37.  These peptides bind to
CC   bacterial lipopolysaccharide (LPS).  These peptides may be used for
CC   treating preventing or minimising septic shock, treating physiological
CC   effects caused by bacterial LPS, treating the effects of LPS on systemic
CC   vascular resistance, cardiac output, tissue perfusion and white blood
CC   cell count, and mediating or preventing the effects of a LPS-induced
CC   endotoxic response in a test animal. These peptides show a strong
CC   antimicrobial activity.  The peptides has enhanced bactericidal activity
CC   compared to full length CAP37 and are active against gram positive and
CC   gram negative bacteria.  AAW15418 is antimicrobial at concentrations of
CC   1.25-7.5 x 10-5 and has maximum activity between pH 5 and pH 5.5.
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 18;  Length 25;
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 5, 2002, 14:48:46 ; Search time 23.96 Seconds
(without alignments)
77.288 Million cell updates/sec

Title: US-09-657-276-1032
Perfect score: 138
Sequence: 1 NQGRHFCGGALIHARFVMTAASCFQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101:*
1: /SIDS8/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	138	100.0	25	14	AAR41934	Peptide fragment o
2	138	100.0	25	18	AAW26803	Human Cationic Ant
3	138	100.0	25	18	AAW15418	CAP37 fragment, 20
4	138	100.0	25	20	AAW97174	Peptide derived fo
5	138	100.0	25	21	AAB28489	Human lipopolysacc
6	138	100.0	25	21	AAB02021	CAP37 peptide anal
7	138	100.0	25	21	AAB09887	Cationic antimicro
8	138	100.0	25	21	AAB17394	CAP37 mimetic/LPS
9	138	100.0	25	22	AAB91856	Antimicrobial pept
10	138	100.0	197	10	AAP91931	Human heparin bind
11	138	100.0	221	16	AAR84663	Human heparin-bind
12	138	100.0	221	17	AAW04875	Heparin binding pr
13	138	100.0	222	14	AAR41935	Recombinant cation
14	138	100.0	222	20	AAW73210	CAP37 protein. Ho
15	138	100.0	225	20	AAY21551	Human heparin-bind
16	138	100.0	225	20	AAW88118	Mature human hepar
17	138	100.0	225	20	AAW88362	Human mature hepar
18	138	100.0	225	21	AAY71876	Human mature hepar
19	138	100.0	225	21	AAY71881	Human heparin bind
20	138	100.0	225	21	AAY71883	Human heparin bind
21	138	100.0	225	22	AAY71891	Human mature hepar
22	138	100.0	226	12	AAR10669	Cationic granule p
23	138	100.0	232	20	AAY21550	Human heparin-bind
24	138	100.0	232	20	AAW88120	Human heparin-bind
25	138	100.0	232	20	AAW88364	Human pro-heparin
26	138	100.0	232	21	AAY71877	Human pre-heparin
27	138	100.0	244	22	AAY71893	Human preheparin-b
28	138	100.0	251	12	AAR10668	Cationic granule p
29	138	100.0	251	16	AAR84666	Human heparin-bind
30	138	100.0	251	20	AAW88121	Complete human hep
31	138	100.0	251	20	AAW88365	Human prepro-hepar
32	138	100.0	251	21	AAY71878	Human pre-pro hepa
33	138	100.0	251	22	AAY71894	Human preprohepari
34	138	100.0	274	16	AAR84665	Human heparin-bind
35	137	99.3	25	21	AAB02058	CAP37 peptide anal
36	135	97.8	25	21	AAB02056	CAP37 peptide anal
37	135	97.8	25	21	AAB02057	CAP37 peptide anal
38	135	97.8	25	21	AAB02060	CAP37 peptide anal
39	135	97.8	25	21	AAB02061	CAP37 peptide anal
40	135	97.8	25	21	AAB02062	CAP37 peptide anal
41	134	97.1	25	21	AAB02031	CAP37 peptide anal
42	134	97.1	25	21	AAB02059	CAP37 peptide anal
43	133	96.4	25	21	AAB02029	CAP37 peptide anal
44	133	96.4	25	21	AAB02030	CAP37 peptide anal
45	132	95.7	25	21	AAB02047	CAP37 peptide anal

ALIGNMENTS

RESULT 1
AAR41934
ID AAR41934 standard; peptide; 25 AA.
XX
AC AAR41934;
XX
DT 19-APR-1994 (first entry)
XX
DE Peptide fragment of mature CAP37 protein.
XX
KW Peptide; CAP37; cancer; monocytes; chemotaxis; wound healing;
KW antibacterial; infection; neoplastic disease; parasitic disease;
KW endotoxaemia.
XX
OS Homo sapiens.
XX
PN WO9319087-A.
XX
PD 30-SEP-1993.
XX
PF 19-MAR-1993; 93WO-US02580.
XX
PR 19-MAR-1992; 92US-0855417.
XX
PA (UYEM-) UNIV EMORY SCHOOL MEDICINE.
XX
PI Pereira HA, Spitznagel JK;
XX
DR WPI; 1993-320680/40.
XX
PT Peptide fragments of CAP37 protein - with chemotactic, antibiotic
PT and lipo:polysaccharide-binding activities
XX
PS Claim 9; Page 72; 109pp; English.

```
  Best Local Similarity   100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 nqgrhfcggaliharfvmtaascfq 25


RESULT   4
AAW97174
ID   AAW97174 standard; peptide; 25 AA.
XX
AC   AAW97174;
XX
DT   30-APR-1999  (first entry)
XX
DE   Peptide derived form residues 20-44 of mature CAP37.
XX
KW   CAP37; Cationic Antimicrobial protein antibiotic; human;
KW   antimicrobial; endotoxin; tumour necrosis factor-alpha;
KW   TNF-alpha; endotoxic shock; septic shock; gram negative bacteria.
XX
OS   Homo sapiens.
XX
PN   US5877151-A.
XX
PD   02-MAR-1999.
XX
PF   21-APR-1997;   97US-0840519.
XX
PR   07-JUN-1995;   95US-0482328.
PR   05-JUL-1989;   89US-0375739.
PR   25-JUN-1990;   90US-0543151.
PR   19-MAR-1992;   92US-0855417.
PR   30-OCT-1992;   92US-0969931.
PR   29-APR-1994;   94US-0235399.
PR   21-APR-1997;   97US-0840519.
XX
PA   (OKLA ) UNIV OKLAHOMA STATE.
XX
PI   Pereira HA;
XX
DR   WPI; 1999-189693/16.
XX
PT   Inhibiting tumor necrosis factor alpha production - using peptides
PT   derived from cationic antimicrobial protein CAP37, e.g. to treat
PT   sepsis
XX
PS   Claim 1; Column 5; 29pp; English.
XX
CC   The present sequence represents a peptide derived from residues 20-44 of
CC   a protein called CAP37 (Cationic Antimicrobial protein, having a
CC   molecular weight of 37 kD). CAP37 has antibiotic activity and is a
CC   chemoattractant for monocytes, and can bind heparin, endotoxin and
CC   lipopolysaccharide. CAP367 can be is present in the granules of
CC   human neutrophils. The present peptide also has strong antimicrobial
CC   activity, and can be used to reduce endotoxin-induced production of
CC   tumour necrosis factor (TNF)-alpha. The peptide is useful for treating
CC   ongoing endotoxic (septic) shock and to prophylactically treat an
CC   individual who may have a risk of septic shock prior to a surgical
CC   procedure such as bowel or bladder surgery or surgical manipulation
CC   of other organs where gram negative bacteria normally reside and could
CC   enter the bloodstream.
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 20;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 nqgrhfcggaliharfvmtaascfq 25


RESULT   5
AAB28489
ID   AAB28489 standard; peptide; 25 AA.
XX
AC   AAB28489;
XX
DT   02-FEB-2001  (first entry)
XX
DE   Human lipopolysaccharide binding protein #1.
XX
KW   Human; lipopolysaccharide binding protein; lipopolysaccharide analyte;
KW   LPSA; immunoassay; gram-negative bacterium detection; Escherichia coli;
KW   Salmonella; Chlamydia; infection.
XX
OS   Homo sapiens.
XX
PN   WO200060354-A1.
XX
PD   12-OCT-2000.
XX
PF   03-APR-2000; 2000WO-EP02869.
XX
PR   07-APR-1999;   99EP-0302711.
XX
PA   (UNIL ) UNILEVER PLC.
PA   (UNIL ) UNILEVER NV.
PA   (HIND-) HINDUSTAN LEVER LTD.
XX
PI   Badley RA,  Hughes G,  Zak KW;
XX
DR   WPI; 2000-679390/66.
XX
PT   Immunoassay for detecting lipopolysaccharides (LPS) from gram negative
PT   bacteria in a sample, comprises use of LPS-binding protein and an
PT   antibody to LPS as first or second binding reagents -
XX
PS   Claim 12; Page 13; 40pp; English.
XX
CC   The present sequence is lipopolysaccharide binding protein which may be
CC   used for detecting the presence of a lipopolysaccharide analyte (LPSA) in
CC   a sample. Anti-LPSA antibody is also used to detect the LPSA. The
CC   immunoassay is useful for detecting the presence of LPSA derived from the
CC   cell membrane of a gram-negative bacterium such as Escherichia coli,
CC   Salmonella and Chlamydia. It provides for assay devices to be
CC   constructed which are capable of accurate and reproducible test
CC   results, ideally suited for commercial markets such as the clinical or
CC   home-testing markets. Furthermore, such immunoassays can be performed
CC   quickly and simply, without using complex extracts or multiple assay
CC   steps.
XX
SQ   Sequence   25 AA;


  Query Match           100.0%;  Score 138;  DB 21;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 7.2e-10;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db       1 nqgrhfcggaliharfvmtaascfq 25


RESULT   6
AAB02021
ID   AAB02021 standard; peptide; 25 AA.
XX
AC   AAB02021;
XX
DT   03-JAN-2001  (first entry)
XX
```

XX
PR 23-OCT-1998; 98US-0105371.
PR 22-OCT-1999; 99US-0428082.
XX
PA (AMGE-) AMGEN INC.
XX
PI Feige U, Liu C, Cheetham J, Boone TC;
XX
DR WPI; 2000-350702/30.
XX
PT Novel composition of matter comprising an Fc domain and
PT pharmacologically active peptides, useful for treating cancer and
PT autoimmune diseases -
XX
PS Claim 39; Page 370; 608pp; English.
XX
CC The present invention describes composition of matter (I) comprising an
CC Fc domain, pharmacologically active peptides, and linkers. Where (I) is:
CC (X1)a-F1-(X2)b, where: F1 = an Fc domain; X1 and X2 = are each
CC independently selected from -(L1)c-P1, -(L1)c-P1-(L2)d-P2,
CC -(L1)c-P1-(L2)d-P2-(L3)e-P^3, or -(L1)c-P1-(L2)d-P2-(L3)e-P3-(L4)f-P4
CC where P1, P2, P3, and P4 = are each independently sequences of
CC pharmacologically active peptides; L1, L2, L3, and L4 = are each
CC independently linkers; and a, b, c, d, e, and f = are each independently
CC 0 or 1, provided that at least 1 of a and b is 1. The composition can
CC have cytostatic, antiasthmatic, thrombolytic and immunosuppressive
CC activities. DNAs, vectors and host cells from the present invention can
CC be used for producing pharmaceutical compositions. The compositions are
CC useful for treating cancer, asthma, thrombosis, or autoimmune diseases.
CC The use of an Fc domain (rather than a Fab domain) can provide a longer
CC half-life or incorporate functions such as Fc receptor binding, protein
CC A binding, complement fixation, and possibly placental transfer. AAA69443
CC to AAA69526 and AAB16955 to AAB18003 represent nucleotide and amino acid
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 138; DB 21; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
||||||||||||||||||||||||||
Db 1 nqgrhfcggaliharfvmtaascfq 25

RESULT 9
AAB91856
ID AAB91856 standard; Peptide; 25 AA.
XX
AC AAB91856;
XX
DT 22-JUN-2001 (first entry)
XX
DE Antimicrobial peptide SEQ ID NO:1032.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
XX
PS Disclosure; Page 533; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 138; DB 22; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.2e-10;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
||||||||||||||||||||||||||
Db 1 nqgrhfcggaliharfvmtaascfq 25

RESULT 10
AAP91931
ID AAP91931 standard; protein; 197 AA.
XX
AC AAP91931;
XX
DT 06-FEB-1990 (first entry)
XX
DE Human heparin binding protein.
XX
KW Heparin binding protein; angiogenic.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 94
FT Modified-site 98
FT Modified-site 131
FT Region 74-75
FT Region 153
FT Region 162
FT Region 86
FT Region 122
FT Region 167
XX
PN WO8908666-A.
XX
PD 21-SEP-1989.
XX
PF 17-MAR-1989; 89WO-DK00059.

XX
PT Conjugate of heparin binding protein with ceramide analogue - used
PT to treat conditions involving stress injury to cells e.g.
PT inflammation, viral infection, sepsis and septic shock
XX
PS Disclosure; Page 22-23; 39pp; English.
XX
CC Heparin binding protein conjugated to a ceramide analogue may be
CC used in the prevention or treatment of inflammation, viral infection,
CC ischaemic reperfusion syndrome, sepsis, septic shock, disseminated
CC intravascular coagulation or for stimulating a patient's immune
CC system. When the conjugate is contacted with living cells, the
CC lipid containing ceramide analogue activates a ceramide activated
CC protein phosphatase which results in down-regulation of cellular
CC metabolism.
XX
SQ Sequence 221 AA;

Query Match 100.0%; Score 138; DB 17; Length 221;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
||||||||||||||||||||||||||
Db 20 nqgrhfcggaliharfvmtaascfq 44

RESULT 13
AAR41935
ID AAR41935 standard; peptide; 222 AA.
XX
AC AAR41935;
XX
DT 19-APR-1994 (first entry)
XX
DE Recombinant cationic granule protein (CAP37).
XX
KW Peptide; CAP37; cancer; monocytes; chemotaxis; wound healing;
KW antibacterial; infection; neoplastic disease; parasitic disease;
KW endotoxaemia; SLE; herpes simplex.
XX
OS Homo sapiens.
XX
PN WO9319087-A.
XX
PD 30-SEP-1993.
XX
PF 19-MAR-1993; 93WO-US02580.
XX
PR 19-MAR-1992; 92US-0855417.
XX
PA (UYEM-) UNIV EMORY SCHOOL MEDICINE.
XX
PI Pereira HA, Spitznagel JK;
XX
DR WPI; 1993-320680/40.
XX
PT Peptide fragments of CAP37 protein - with chemotactic, antibiotic
PT and lipo:polysaccharide-binding activities
XX
PS Disclosure; Page 73; 109pp; English.
XX
CC CAP37 has chemotactic activity for monocytes, is bactericidal and
CC is capable of binding bacterial lipopolysaccharide.
CC It is especially useful for treating wounds or for dental
CC applications. Other uses of CAP37 or its peptides include the
CC treatment of cancer and any disease involving monocyte localisation
CC including neoplastic diseases, parasitic diseases and rheumatoid
CC arthritis. CAP37 may also be used to treat diseases involving
CC defects of monocyte chemotaxis such as chronic mucocutaneous
CC candidiasis, SLE and herpes simplex. Specific CAP37 peptides may
CC be useful for treating infection due to antibacterial activity or
CC for wound treatment because they possess antibiotic or
CC lipopolysaccharide-binding activity. CAP37 and constituent
CC peptides may also be used to treat endotoxaemia. The recombinant
CC CAP37 molecule can have N-terminal extensions added to it, these
CC being either a single methionine residue or the sequence described
CC in file AAR41936.
XX
SQ Sequence 222 AA;

Query Match 100.0%; Score 138; DB 14; Length 222;
Best Local Similarity 100.0%; Pred. No. 4.3e-09;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
||||||||||||||||||||||||||
Db 20 nqgrhfcggaliharfvmtaascfq 44

RESULT 14
AAW73210
ID AAW73210 standard; peptide; 222 AA.
XX
AC AAW73210;
XX
DT 22-JAN-1999 (first entry)
XX
DE CAP37 protein.
XX
KW Defensin; T-cell chemotaxis; therapy; immune response induction; tumour;
KW antigenic agent; infection; hyperproliferative disorder; CAP37;
KW immunodeficiency.
XX
OS Homo sapiens.
XX
PN US5837247-A.
XX
PD 17-NOV-1998.
XX
PF 16-JUN-1995; 95US-0491204.
XX
PR 16-JUN-1995; 95US-0491204.
XX
PA (USSH) US PUBLIC HEALTH SERVICE NAT INST HEALTH.
XX
PI Chertov O, Michiel D, Murphy WJ, Oppenhelm JJ, Taub DD;
PI Wang JM, Xu L;
XX
DR WPI; 1999-023377/02.
XX
PT Use of defensin proteins - for inducing or stimulating T-cell
PT chemotaxis used for treating infections, tumours, immunodeficiencies
PT or autoimmune conditions
XX
PS Disclosure; Column 25-28; 20pp; English.
XX
CC This sequence represents the CAP37 protein, and can be used
CC in the method of the invention. The method is for inducing or
CC stimulating T-cell chemotaxis in a subject, comprising administering to
CC the subject a pharmaceutical composition comprising a defensin protein.
CC The method can be used for inducing an immune response in a subject to an
CC antigenic agent such as tumour, infectious agent or other diseases
CC tissue. The method can also be used to treat e.g. bacterial, viral,
CC fungal and other infections, tumours and other hyperproliferative
CC disorders, immunodeficiencies, diseases susceptible to treatment by
CC administration of a therapeutic vaccine and autoimmune conditions. The
CC peptides can also be used to treat inflammation or autoimmune
CC conditions.
XX
SQ Sequence 222 AA;

Query Match 100.0%; Score 138; DB 20; Length 222;

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 5, 2002, 14:57:46 ; Search time 12.8 Seconds
(without alignments)
148.778 Million cell updates/sec

Title: US-09-657-276-1032
Perfect score: 138
Sequence: 1 NQGRHFCGGALIHARFVMTAASCFQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	138	100.0	251	1	TRHUAZ	azurocidin precurs
2	103	74.6	219	1	TRPGAZ	azurocidin - pig
3	93	67.4	256	1	PRHU3	proteinase 3 (EC 3
4	85	61.6	231	1	TRPCTR	trypsin (EC 3.4.21
5	85	61.6	248	2	S55066	trypsin (EC 3.4.21
6	85	61.6	258	2	I56220	tryptase 2 - rat
7	85	61.6	270	2	S56160	mast cell tryptase
8	84	60.9	229	1	TRBOTR	trypsin (EC 3.4.21
9	84	60.9	231	2	S31778	trypsin (EC 3.4.21
10	84	60.9	246	1	TRDGC	trypsin (EC 3.4.21
11	84	60.9	247	2	S13813	trypsin (EC 3.4.21
12	83	60.1	247	2	A27547	trypsin (EC 3.4.21
13	83	60.1	248	2	S55067	trypsin (EC 3.4.21
14	83	60.1	711	1	A47136	macrophage-stimula
15	82	59.4	455	2	A61545	plasmin (EC 3.4.21
16	81	58.7	246	1	TRRT1	trypsin (EC 3.4.21
17	81	58.7	246	1	TRRT2	trypsin (EC 3.4.21
18	81	58.7	247	2	S12764	trypsin (EC 3.4.21
19	81	58.7	259	2	I38363	trypsin (EC 3.4.21
20	81	58.7	304	2	S33496	trypsin (EC 3.4.21
21	80	58.0	230	2	I48685	mast cell proteina
22	80	58.0	242	2	S31775	trypsin (EC 3.4.21
23	80	58.0	242	2	S49489	trypsin (EC 3.4.21
24	80	58.0	242	2	S31776	trypsin (EC 3.4.21
25	80	58.0	246	2	B25528	trypsin (EC 3.4.21
26	80	58.0	273	2	A47246	tryptase (EC 3.4.2
27	80	58.0	274	2	A45754	tryptase (EC 3.4.2
28	80	58.0	274	2	JC4171	tryptase (EC 3.4.2
29	80	58.0	275	2	A35863	tryptase (EC 3.4.2
30	80	58.0	275	2	B35863	tryptase (EC 3.4.2
31	80	58.0	275	2	C35863	tryptase (EC 3.4.2
32	80	58.0	276	2	A38654	mast cell proteina
33	79.5	57.6	558	2	JC5878	plasma hyaluronan-
34	79.5	57.6	560	1	JC4795	plasma hyaluronan-
35	79	57.2	196	2	T08808	hypothetical prote
36	79	57.2	254	1	TRWV3Y	trypsin-like prote
37	79	57.2	269	2	B32410	mastocytoma protei
38	79	57.2	275	2	A32410	tryptase (EC 3.4.2
39	79	57.2	1420	2	A32869	apolipoprotein(a)
40	78	56.5	229	1	TRDFS	trypsin (EC 3.4.21
41	78	56.5	237	2	S68702	tryptase (EC 3.4.2
42	78	56.5	243	2	A35871	trypsin (EC 3.4.21
43	78	56.5	247	1	A25852	trypsin (EC 3.4.21
44	78	56.5	247	1	TRDG	trypsin (EC 3.4.21
45	78	56.5	249	2	A55634	granzyme M (EC 3.4

ALIGNMENTS

RESULT 1
TRHUAZ
azurocidin precursor [validated] - human
N;Alternate names: cationic antimicrobial protein CAP37; heparin-binding protein; ne
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 30-Sep-1993 #text_change 08-Dec-2000
C;Accession: A46268; A46455; S16450; S18520; S12881; S15445; S14738; B33913; A60708;
R;Zimmer, M.; Medcalf, R.L.; Fink, T.M.; Mattmann, C.; Lichter, P.; Jenne, D.E.
Proc. Natl. Acad. Sci. U.S.A. 89, 8215-8219, 1992
A;Title: Three human elastase-like genes coordinately expressed in the myelomonocyte
A;Reference number: A46268; MUID:92390417
A;Accession: A46268
A;Molecule type: DNA
A;Residues: 1-251 <ZIM>
A;Cross-references: GB:M96326; NID:g179301; PIDN:AAB59353.1; PID:g179302
A;Note: sequence extracted from NCBI backbone (NCBIN:112883, NCBIN:112891, NCBIN:112
R;Morgan, J.G.; Sukiennicki, T.; Pereira, H.A.; Spitznagel, J.K.; Guerra, M.E.; Larr
J. Immunol. 147, 3210-3214, 1991
A;Title: Cloning of the cDNA for the serine protease homolog CAP37/azurocidin, a mic
A;Reference number: A46455; MUID:92013155
A;Accession: A46455
A;Molecule type: mRNA
A;Residues: 1-251 <MOR>
A;Cross-references: GB:M96326; NID:g179301; PIDN:AAB59353.1; PID:g179302
A;Note: sequence extracted from NCBI backbone (NCBIN:60386, NCBIP:60395)
R;Almeida, R.P.; Melchior, M.; Campanelli, D.; Nathan, C.; Gabay, J.E.
Biochem. Biophys. Res. Commun. 177, 688-695, 1991
A;Title: Complementary DNA sequence of human neutrophil azurocidin, an antibiotic wi
A;Reference number: S16450; MUID:91264832
A;Accession: S16450
A;Molecule type: mRNA
A;Residues: 3-251 <ALM>
A;Cross-references: EMBL:X58794; NID:g28976; PIDN:CAA41601.1; PID:g28977
A;Accession: S18520
A;Molecule type: protein
A;Residues: 47-51,'S',53-59,'T',61-62,'V',64-87 <AL2>
R;Pohl, J.; Pereira, H.A.; Martin, N.M.; Spitznagel, J.K.
FEBS Lett. 272, 200-204, 1990
A;Title: Amino acid sequence of CAP37, a human neutrophil granule-derived antibacter
A;Reference number: S12881; MUID:91032128
A;Accession: S12881
A;Molecule type: protein
A;Residues: 27-248 <POH>
R;Flodgaard, H.; Ostergaard, E.; Bayne, S.; Svendsen, A.; Thomsen, J.; Engels, M.; F
Eur. J. Biochem. 197, 535-547, 1991
A;Title: Covalent structure of two novel neutrophile leucocyte-derived proteins of p
A;Reference number: S15393; MUID:91224149
A;Accession: S15445
A;Molecule type: protein
A;Residues: 27-129,'N',131-247 <FLO>
A;Note: the sequence from Fig. 7 is inconsistent with that from Fig. 6 in having 19(

A;Molecule type: DNA
A;Residues: 1-118,'V',120-134,'AT',137-256 <ZIM>
A;Note: sequence extracted from NCBI backbone (NCBIN:112898, NCBIN:112900, NCBIN:112902,
R;Labbaye, C.; Musette, P.; Cayre, Y.E.
Proc. Natl. Acad. Sci. U.S.A. 88, 9253-9256, 1991
A;Title: Wegener autoantigen and myeloblastin are encoded by a single mRNA.
A;Reference number: A43983; MUID:92021028
A;Accession: A43983
A;Molecule type: mRNA
A;Residues: 1-69,'P',71-256 <LA2>
A;Cross-references: GB:M75154; NID:g187398; PIDN:AAA59558.1; PID:g187399
R;Campanelli, D.; Melchior, M.; Fu, Y.; Nakata, M.; Shuman, H.; Nathan, C.; Gabay, J.E.
J. Exp. Med. 172, 1709-1715, 1990
A;Title: Cloning of cDNA for proteinase 3: a serine protease, antibiotic, and autoantige
A;Reference number: JH0331; MUID:91079774
A;Accession: JH0331
A;Molecule type: mRNA
A;Residues: 'R',3-118,'V',120-134,'AT',137-254,'P' <CAM>
A;Cross-references: GB:X55668; NID:g35687; PIDN:CAA39203.1; PID:g1335280
A;Note: part of this sequence, including the amino end of the mature protein, was confir
R;Bories, D.; Raynal, M.C.; Solomon, D.H.; Darzynkiewicz, Z.; Cayre, Y.E.
Cell 59, 959-968, 1989
A;Title: Down-regulation of a serine protease, myeloblastin, causes growth arrest and di
A;Reference number: A33751; MUID:90090622
A;Accession: A33751
A;Molecule type: mRNA
A;Residues: 42-256 <BOR>
A;Cross-references: GB:M29142; NID:g188983; PIDN:AAA36342.1; PID:g188984
A;Note: the authors translated the codon GGG for residue 49 as Glu, GGC for residue 52 a
R;Jenne, D.E.; Tschopp, J.; Luedemann, J.; Utecht, B.; Gross, W.L.
Nature 346, 520, 1990
A;Title: Wegener's autoantigen decoded.
A;Reference number: S11091; MUID:90332035
A;Accession: S11091
A;Molecule type: mRNA
A;Residues: 20-56 <JEN>
R;Musette, P.; Labbaye, C.; Dorner, M.H.; Cayre, Y.E.; Casanova, J.L.; Kourilsky, P.
Blood 77, 1398-1399, 1991
A;Title: Wegener's autoantigen and leukemia.
A;Reference number: A61176; MUID:91159650
A;Accession: A61176
A;Molecule type: mRNA
A;Residues: 1-42 <MUS>
A;Cross-references: EMBL:X56606; NID:g35189; PIDN:CAA39943.1; PID:g35190
R;Goldschmeding, R.; Dolman, K.M.; Van Den Ende, M.E.; Van Der Meer-Gerritsen, C.H.; Son
APMIS 19(Suppl.), 26-27, 1990
A;Title: The relation of 29 kD C-ANCA antigen to proteinase 3.
A;Reference number: A60148; MUID:91136884
A;Accession: A60148
A;Molecule type: protein
A;Residues: 28-48 <GOL>
R;Rao, N.V.; Wehner, N.G.; Marshall, B.C.; Gray, W.R.; Gray, B.H.; Hoidal, J.R.
J. Biol. Chem. 266; 9540-9548, 1991
A;Title: Characterization of proteinase-3 (PR-3), a neutrophil serine proteinase. Struct
A;Reference number: A43982; MUID:91236723
A;Accession: A43982
A;Molecule type: protein
A;Residues: 28-61,'X',63,'D',65-67;228-244 <RAO>
R;Wilde, C.G.; Snable, J.L.; Griffith, J.E.; Scott, R.W.
J. Biol. Chem. 265, 2038-2041, 1990
A;Title: Characterization of two azurophil granule proteases with active-site homology t
A;Reference number: A43981; MUID:90130450
A;Accession: A43981
A;Molecule type: protein
A;Residues: 28-45,'E',47;196-208,'X',210-215,'X',217-219 <WIL>
R;Gabay, J.E.; Scott, R.W.; Campanelli, D.; Griffith, J.; Wilde, C.; Marra, M.N.; Seeger
Proc. Natl. Acad. Sci. U.S.A. 86, 5610-5614, 1989
A;Title: Antibiotic proteins of human polymorphonuclear leukocytes.
A;Reference number: A33913; MUID:89315847
A;Accession: C33913
A;Molecule type: protein
A;Residues: 28-45,'E',47 <GAB>
R;Niles, J.L.; McCluskey, R.T.; Ahmad, M.F.; Arnaout, M.A.
Blood 74, 1888-1893, 1989
A;Title: Wegener's granulomatosis autoantigen is a novel neutrophil serine proteinas
A;Reference number: A60481; MUID:90028708
A;Accession: A60481
A;Molecule type: protein
A;Residues: 28-38,'X',40-47 <NI2>
R;Ohlsson, K.; Linder, C.; Rosengren, M.
Biol. Chem. Hoppe-Seyler 371, 549-555, 1990
A;Title: Monoclonal antibodies specific for neutrophil proteinase 4. Production and
A;Reference number: S10605; MUID:91025622
A;Accession: S10605
A;Molecule type: protein
A;Residues: 28-52 <OHL>
R;Luedemann, J.; Utecht, B.; Gross, W.L.
J. Exp. Med. 171, 357-362, 1990
A;Title: Anti-neutrophil cytoplasm antibodies in Wegener's granulomatosis recognize
A;Reference number: PL0230; MUID:90111630
A;Accession: PL0230
A;Molecule type: protein
A;Residues: 28-37,'I',39-40,'I',41-43 <LUE>
C;Comment: This polymorphonuclear leukocyte serine protease from azurophilic granule
C;Genetics:
A;Gene: GDB:PRTN3
A;Cross-references: GDB:126876; OMIM:177020
A;Map position: 19p13.3-19p13.3
A;Introns: 21/1; 76/2; 123/3; 200/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; polymorphonuclear leukocyte; serine proteinase
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-27/Domain: propeptide #status predicted <PRO>
F;28-256/Product: proteinase 3 #status experimental <MAT>
F;28-243/Domain: trypsin homology <TRY>
F;56-72,152-209,182-188,199-224/Disulfide bonds: #status predicted
F;71,118,203/Active site: His, Asp, Ser #status predicted
F;129,174/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 67.4%; Score 93; DB 1; Length 256;
Best Local Similarity 64.0%; Pred. No. 1.6e-06;
Matches 16; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

Qy 1 NQGRHFCGGALIHARFVMTAASCFQ 25
| | ||||| ||| ||:||| | :
Db 50 NPGSHFCGGTLIHPSFVLTAAHCLR 74

RESULT 4
TRPGTR
trypsin (EC 3.4.21.4) precursor - pig (tentative sequence)
N;Contains: trypsinogen
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 24-Apr-1984 #sequence_revision 24-Apr-1984 #text_change 31-Mar-2000
C;Accession: A90641; A90368; A00947
R;Charles, M.; Rovery, M.; Guidoni, A.; Desnuelle, P.
Biochim. Biophys. Acta 69, 115-129, 1963
A;Title: Su le trypsinogene et la trypsine de porc.
A;Reference number: A90641
A;Accession: A90641
A;Molecule type: protein
A;Residues: 1-10 <CHA>
R;Hermodson, M.A.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 12, 3146-3153, 1973
A;Title: Determination of the amino acid sequence of porcine trypsin by sequenator a
A;Reference number: A90368; MUID:73258692
A;Accession: A90368
A;Molecule type: protein
A;Residues: 9-231 <HER>
A;Note: at position 20, Ile and Val occur alternatively
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; polymorphism; protein digestion; serine proteinase;
F;1-231/Product: trypsinogen #status experimental <ZYM>
F;1-8/Domain: activation peptide #status experimental <APT>
F;9-231/Product: trypsin #status experimental <MAT>

```
F;7-222/Domain: trypsin homology <TRY>
F;7-131,132-229/Product: alpha-trypsin #status experimental <MPT>
F;6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental
F;13-143,31-47,115-216,122-189,154-168,179-203/Disulfide bonds: #status experimental
F;46,90,183/Active site: His, Asp, Ser #status experimental
F;58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental
F;131-132/Cleavage site: Lys-Ser (autolytic) #status experimental


  Query Match             60.9%;  Score 84;  DB 1;  Length 229;
  Best Local Similarity   52.0%;  Pred. No. 3.2e-05;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      25 NSGYHFCGGSLINSQWVVSAAHCYK 49


RESULT  9
S31778
trypsin (EC 3.4.21.4) II precursor - Atlantic salmon (fragment)
C;Species: Salmo salar (Atlantic salmon)
C;Date: 22-Nov-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C;Accession: S66658; S31778
R;Male, R.; Lorens, J.B.; Smalas, A.O.; Torrissen, K.R.
Eur. J. Biochem. 232, 677-685, 1995
A;Title: Molecular cloning and characterization of anionic and cationic variants of tryp
A;Reference number: S66657; MUID:96035908
A;Accession: S66658
A;Molecule type: mRNA
A;Residues: 1-231 <MAL>
A;Cross-references: EMBL:X70073; NID:g64385; PIDN:CAA49678.1; PID:g64386
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;1-4/Domain: signal sequence (fragment) #status predicted <SIG>
F;5-9/Domain: activation peptide #status predicted <APT>
F;10-231/Product: trypsin II #status predicted <MAT>
F;10-224/Domain: trypsin homology <TRY>
F;16-145,34-50,118-218,125-191,156-170,181-205/Disulfide bonds: #status predicted
F;49,93,185/Active site: His, Asp, Ser #status predicted


  Query Match             60.9%;  Score 84;  DB 2;  Length 231;
  Best Local Similarity   52.0%;  Pred. No. 3.2e-05;
  Matches   13;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:|::  :|:::|| |:|
Db      28 NSGYHFCGGSLVNENWVVSAAHCYQ 52


RESULT  10
TRDGC
trypsin (EC 3.4.21.4) precursor, cationic - dog
N;Alternate names: cationic trypsinogen
C;Species: Canis lupus familiaris (dog)
C;Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 18-Jun-1999
C;Accession: B26273
R;Pinsky, S.D.; LaForge, K.S.; Scheele, G.
Mol. Cell. Biol. 5, 2669-2676, 1985
A;Title: Differential regulation of trypsinogen mRNA translation: full-length mRNA sequ
A;Reference number: A26273; MUID:86284628
A;Accession: B26273
A;Molecule type: mRNA
A;Residues: 1-246 <PIN>
A;Cross-references: GB:M11590; NID:g164096; PIDN:AAA30900.1; PID:g164097
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-23/Domain: activation peptide #status predicted <APT>
F;24-246/Product: trypsin, cationic #status predicted <ENZ>
F;24-239/Domain: trypsin homology <TRY>
F;30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F;63,107,200/Active site: His, Asp, Ser #status predicted
F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted


  Query Match             60.9%;  Score 84;  DB 1;  Length 246;
  Best Local Similarity   52.0%;  Pred. No. 3.4e-05;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      42 NSGYHFCGGSLINSQWVVSAAHCYK 66


RESULT  11
S13813
trypsin (EC 3.4.21.4) - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 02-Dec-1993 #sequence_revision 03-Aug-1995 #text_change 22-Jun-1999
C;Accession: S13813
R;le Huerou, I.; Wicker, C.; Guilloteau, P.; Toullec, R.; Puigserver, A.
Eur. J. Biochem. 193, 767-773, 1990
A;Title: Isolation and nucleotide sequence of cDNA clone for bovine pancreatic anion
A;Reference number: S13813; MUID:91065383
A;Accession: S13813
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-247 <HUE>
A;Cross-references: EMBL:X54703; NID:g829; PIDN:CAA38513.1; PID:g830
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; protein digestion; serine proteinase
F;24-239/Domain: trypsin homology <TRY>
F;63,107,200/Active site: His, Asp, Ser #status predicted


  Query Match             60.9%;  Score 84;  DB 2;  Length 247;
  Best Local Similarity   56.0%;  Pred. No. 3.4e-05;
  Matches   14;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||: ::|::|| |:|
Db      42 NAGYHFCGGSLINDQWVVSAAHCYQ 66


RESULT  12
A27547
trypsin (EC 3.4.21.4) precursor, cationic - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 22-Jun-1999
C;Accession: A27547
R;Fletcher, T.S.; Alhadeff, M.; Craik, C.S.; Largman, C.
Biochemistry 26, 3081-3086, 1987
A;Title: Isolation and characterization of a cDNA encoding rat cationic trypsinogen.
A;Reference number: A27547; MUID:87271609
A;Accession: A27547
A;Molecule type: mRNA
A;Residues: 1-247 <FLE>
A;Cross-references: GB:M16624; NID:g206498; PIDN:AAA41985.1; PID:g206499
C;Superfamily: trypsin; trypsin homology
C;Keywords: calcium binding; hydrolase; protein digestion; serine proteinase
F;25-240/Domain: trypsin homology <TRY>
F;31-161,49-65,133-234,140-207,172-186/Disulfide bonds: #status predicted
F;64,108,201/Active site: His, Asp, Ser #status predicted
F;76,78,81,86/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted


  Query Match             60.1%;  Score 83;  DB 2;  Length 247;
  Best Local Similarity   52.0%;  Pred. No. 4.8e-05;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      43 NAGYHFCGGSLINSQWVVSAAHCYK 67
```

F;1-33/Domain: activation peptide (fragment) #status experimental <APT>
F;34-117,118-225,226-455/Product: plasmin (fragments) #status experimental <MAT>
F;37-114/Domain: kringle homology <KR4>
F;118-455/Product: miniplasminogen #status experimental <MIN>
F;126-205/Domain: kringle homology <KR5>
F;226-455/Domain: plasmin chain B #status experimental <BCH>
F;226-448/Domain: trypsin homology <TRY>
F;267,310,405/Active site: His, Asp, Ser #status predicted

Query Match 59.4%; Score 82; DB 2; Length 455;
Best Local Similarity 60.9%; Pred. No. 0.00011;
Matches 14; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

```
Qy       3 GRHFCGGALIHARFVMTAASCFQ 25
           ||||||| ||   :|:||| | :
Db     248 GRHFCGGTLISPEWVLTAAHCLE 270
```

Search completed: April 5, 2002, 15:00:25
Job time: 159 sec

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 5, 2002, 14:59:26 ; Search time 10.2 Seconds
(without alignments)
89.865 Million cell updates/sec

Title: US-09-657-276-1032
Perfect score: 138
Sequence: 1 NQGRHFCGGALIHARFVMTAASCFQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	100.0	251	1	CAP7_HUMAN	P20160 homo sapien
2	103	74.6	219	1	CAP7_PIG	P80015 sus scrofa
3	93	67.4	256	1	PRN3_HUMAN	P24158 homo sapien
4	85	61.6	231	1	TRYP_PIG	P00761 sus scrofa
5	85	61.6	248	1	TRY3_CHICK	Q90629 gallus gall
6	85	61.6	258	1	GRAK_RAT	P49864 rattus norv
7	85	61.6	270	1	TRYT_MERUN	P50342 meriones un
8	84	60.9	231	1	TRY2_SALSA	P35032 salmo salar
9	84	60.9	243	1	TRY1_BOVIN	P00760 bos taurus
10	84	60.9	246	1	TRY1_CANFA	P06871 canis famil
11	84	60.9	247	1	TRY2_BOVIN	Q29463 bos taurus
12	83	60.1	244	1	TRY2_XENLA	P70059 xenopus lae
13	83	60.1	247	1	TRY3_RAT	P08426 rattus norv
14	83	60.1	248	1	TRY1_CHICK	Q90627 gallus gall
15	83	60.1	248	1	TRY2_CHICK	Q90628 gallus gall
16	83	60.1	711	1	HGFL_HUMAN	P26927 homo sapien
17	82	59.4	338	1	PLMN_HORSE	P80010 equus cabal
18	81	58.7	246	1	TRY1_RAT	P00762 rattus norv
19	81	58.7	246	1	TRY2_RAT	P00763 rattus norv
20	81	58.7	247	1	TRY3_HUMAN	P15951 homo sapien
21	81	58.7	304	1	TRY4_HUMAN	P35030 homo sapien
22	80	58.0	242	1	TRY1_SALSA	P35031 salmo salar
23	80	58.0	246	1	TRY2_MOUSE	P07146 mus musculu
24	80	58.0	273	1	MCT7_MOUSE	Q02844 mus musculu
25	80	58.0	273	1	MCT7_RAT	P27435 rattus norv
26	80	58.0	274	1	TRYM_RAT	P50343 rattus norv
27	80	58.0	275	1	TRYA_HUMAN	P15157 homo sapien
28	80	58.0	275	1	TRYB_HUMAN	P20231 homo sapien
29	80	58.0	276	1	MCT6_MOUSE	P21845 mus musculu
30	79	57.2	254	1	TRY3_AEDAE	P29786 aedes aegyp
31	79	57.2	263	1	GRAK_MOUSE	O35205 mus musculu
32	79	57.2	269	1	TRYM_CANFA	P19236 canis famil
33	79	57.2	275	1	TRYT_CANFA	P15944 canis famil
34	79	57.2	277	1	KLKD_HUMAN	Q9ukr3 homo sapien
35	79	57.2	324	1	TEST_MOUSE	Q9jhj7 mus musculu
36	79	57.2	333	1	PLMN_CANFA	P80009 canis famil
37	79	57.2	1420	1	APOA_MACMU	P14417 macaca mula
38	78	56.5	229	1	TRYP_SQUAC	P00764 squalus aca
39	78	56.5	243	1	TRY1_XENLA	P19799 xenopus lae
40	78	56.5	247	1	TRY1_HUMAN	P07477 homo sapien
41	78	56.5	247	1	TRY2_CANFA	P06872 canis famil
42	78	56.5	257	1	GRAM_HUMAN	P51124 homo sapien
43	78	56.5	716	1	HGFL_MOUSE	P26928 mus musculu
44	78	56.5	4548	1	APOA_HUMAN	P08519 homo sapien
45	77	55.8	253	1	TRYB_DROME	P35004 drosophila

ALIGNMENTS

RESULT 1
CAP7_HUMAN
ID CAP7_HUMAN STANDARD; PRT; 251 AA.
AC P20160; P80014;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE AZUROCIDIN PRECURSOR (CATIONIC ANTIMICROBIAL PROTEIN CAP37)
DE (HEPARIN-BINDING PROTEIN) (HBP).
GN AZU1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92013155; PubMed=1919011;
RA Morgan J.G., Sukiennicki T., Pereira H.A., Spitznagel J.K.,
RA Guerra M.E., Larrick J.L.;
RT "Cloning of the cDNA for the serine protease homolog
RT CAP37/azurocidin, a microbicidal and chemotactic protein from human
RT granulocytes.";
RL J. Immunol. 147:3210-3214(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92390417; PubMed=1518849;
RA Zimmer M., Medcalf R.L., Fink T.M., Mattmann C., Lichter P.,
RA Jenne D.E.;
RT "Three human elastase-like genes coordinately expressed in the
RT myelomonocyte lineage are organized as a single genetic locus on
RT 19pter.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:8215-8219(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA Burkhart-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA Kobayashi A., Olsen A.S., Carrano A.V.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE OF 3-251 FROM N.A.
RC TISSUE=Neutrophils;
RX MEDLINE=91264832; PubMed=2049091;
RA Almeida R.P., Melchior M., Campanelli D., Nathan C., Gabay J.E.;
RT "Complementary DNA sequence of human neutrophil azurocidin, an
RT antibiotic with extensive homology to serine proteases.";
RL Biochem. Biophys. Res. Commun. 177:688-695(1991).
RN [5]
RP SEQUENCE OF 27-248.
RX MEDLINE=91032128; PubMed=2226832;
RA Pohl J., Pereira H.A., Martin N.M., Spitznagel J.K.;
RT "Amino acid sequence of CAP37, a human neutrophil granule-derived

```
  Best Local Similarity   100.0%;  Pred. No. 1.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||||||||||||
Db      46 NQGRHFCGGALIHARFVMTAASCFQ 70


RESULT   2
CAP7_PIG
ID   CAP7_PIG       STANDARD;      PRT;   219 AA.
AC   P80015;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   AZUROCIDIN (CATIONIC ANTIMICROBIAL PROTEIN CAP37) (HEPARIN-BINDING
DE   PROTEIN) (HBP).
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Neutrophils;
RX   MEDLINE=91224149; PubMed=2026172;
RA   Flodgaard H., Oestergaard E., Bayne S., Svendsen A., Thomsen J.,
RA   Engels M., Wollmer A.;
RT   "Covalent structure of two novel neutrophile leucocyte-derived
RT   proteins of porcine and human origin. Neutrophile elastase homologues
RT   with strong monocyte and fibroblast chemotactic activities.";
RL   Eur. J. Biochem. 197:535-547(1991).
CC   -!- FUNCTION: THIS IS A NEUTROPHIL GRANULE-DERIVED ANTIBACTERIAL AND
CC       MONOCYTE- AND FIBROBLAST-SPECIFIC CHEMOTACTIC GLYCOPROTEIN. BINDS
CC       HEPARIN.
CC   -!- SUBUNIT: MONOMER.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY. ELASTASE SUBFAMILY.
DR   PIR; S15393; TRPGAZ.
DR   HSSP; P20160; 1AE5.
DR   MEROPS; S01.971; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.
DR   PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW   Serine protease homolog; Glycoprotein; Chemotaxis; Antibiotic;
KW   Heparin-binding.
FT   CARBOHYD    113    113       N-LINKED (GLCNAC. . .).
FT   CARBOHYD    144    144       N-LINKED (GLCNAC. . .).
FT   DISULFID     26     42       BY SIMILARITY.
FT   DISULFID    122    179       BY SIMILARITY.
FT   DISULFID    152    158       BY SIMILARITY.
SQ   SEQUENCE   219 AA;  24301 MW;  99129CAD88B0749D CRC64;


  Query Match             74.6%;  Score 103;  DB 1;  Length 219;
  Best Local Similarity   75.0%;  Pred. No. 3.6e-09;
  Matches   18;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCFQ 25
           ||| || |||:| |||:||||||:
Db      21 QGRPFCAGALVHPRFVLTAASCFR 44


RESULT   3
PRN3_HUMAN
ID   PRN3_HUMAN     STANDARD;      PRT;   256 AA.
AC   P24158; P15637;
DT   01-APR-1990 (Rel. 14, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   MYELOBLASTIN PRECURSOR (EC 3.4.21.76) (LEUKOCYTE PROTEINASE 3) (PR-3)
DE   (PR3) (AGP7) (WEGENER'S AUTOANTIGEN) (P29) (C-ANCA ANTIGEN).
GN   PRTN3 OR MBN.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92021028; PubMed=1681549;
RA   Labbaye C., Musette P., Cayre Y.E.;
RT   "Wegener autoantigen and myeloblastin are encoded by a single mRNA.";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:9253-9256(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K.J., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J.,
RA   Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J.,
RA   Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G.,
RA   Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G.,
RA   Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S.,
RA   Kobayashi A., Olsen A.S., Carrano A.V.;
RL   Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE OF 2-256 FROM N.A., AND SEQUENCE OF 48-71 AND 156-181.
RX   MEDLINE=91079774; PubMed=2258701;
RA   Campanelli D., Melchior M., Fu Y., Nakata M., Shuman H., Nathan C.,
RA   Gabay J.E.;
RT   "Cloning of cDNA for proteinase 3: a serine protease, antibiotic, and
RT   autoantigen from human neutrophils.";
RL   J. Exp. Med. 172:1709-1715(1990).
RN   [4]
RP   SEQUENCE OF 1-20 AND 22-256 FROM N.A.
RX   MEDLINE=92390417; PubMed=1518849;
RA   Zimmer M., Medcalf R.L., Fink T.M., Mattmann C., Lichter P.,
RA   Jenne D.E.;
RT   "Three human elastase-like genes coordinately expressed in the
RT   myelomonocyte lineage are organized as a single genetic locus on
RT   19pter.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:8215-8219(1992).
RN   [5]
RP   SEQUENCE OF 42-256 FROM N.A.
RX   MEDLINE=90090622; PubMed=2598267;
RA   Bories D., Raynal M.-C., Solomon D.H., Darzynkiewicz Z., Cayre Y.E.;
RT   "Down-regulation of a serine protease, myeloblastin, causes growth
RT   arrest and differentiation of promyelocytic leukemia cells.";
RL   Cell 59:959-968(1989).
RN   [6]
RP   SEQUENCE OF 28-67 AND 228-244.
RX   MEDLINE=91236723; PubMed=2033050;
RA   Rao N.V., Wehner N.G., Marshall B.C., Gray W.R., Gray B.H.,
RA   Hoidal J.R.;
RT   "Characterization of proteinase-3 (PR-3), a neutrophil serine
RT   proteinase. Structural and functional properties.";
RL   J. Biol. Chem. 266:9540-9548(1991).
RN   [7]
RP   SEQUENCE OF 28-47.
RX   MEDLINE=89315847; PubMed=2501794;
RA   Gabay J.E., Scott R.W., Campanelli D., Griffith J., Wilde C.,
RA   Marra M.N., Seeger M., Nathan C.F.;
RT   "Antibiotic proteins of human polymorphonuclear leukocytes.";
RL   Proc. Natl. Acad. Sci. U.S.A. 86:5610-5614(1989).
RN   [8]
RP   SEQUENCE OF 28-47 AND 196-219.
RX   MEDLINE=90130450; PubMed=2404977;
RA   Wilde C.G., Snable J.L., Griffith J.E., Scott R.W.;
RT   "Characterization of two azurphil granule proteases with active-site
RT   homology to neutrophil elastase.";
RL   J. Biol. Chem. 265:2038-2041(1990).
RN   [9]
RP   SEQUENCE OF 28-48, AND IDENTITY OF WEGENER'S AUTOANTIGEN WITH PR-3.
```

```
RL   J. Biol. Chem. 272:19931-19937(1997).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (2.03 ANGSTROMS) IN COMPLEX WITH LDTI.
RX   MEDLINE=98046095; PubMed=9384562;
RA   di Marco S., Priestle J.P.;
RT   "Structure of the complex of leech-derived tryptase inhibitor (LDTI)
RT   with trypsin and modeling of the LDTI-tryptase system.";
RL   Structure 5:1465-1474(1997).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
DR   PIR; A00947; TRPGTR.
DR   PDB; 1MCT; 31-JAN-94.
DR   PDB; 1AKS; 12-FEB-97.
DR   PDB; 1EPT; 07-FEB-95.
DR   PDB; 1TFX; 21-JAN-98.
DR   PDB; 1LDT; 20-MAY-98.
DR   PDB; 1AN1; 01-JUL-98.
DR   PDB; 1AVW; 18-NOV-98.
DR   PDB; 1AVX; 18-NOV-98.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen;
KW   3D-structure.
FT   PROPEP        1      8       ACTIVATION PEPTIDE.
FT   CHAIN         9    231       TRYPSIN.
FT   ACT_SITE     48     48       CHARGE RELAY SYSTEM.
FT   ACT_SITE     92     92       CHARGE RELAY SYSTEM.
FT   ACT_SITE    185    185       CHARGE RELAY SYSTEM.
FT   DISULFID     15    145
FT   DISULFID     33     49
FT   DISULFID    117    218
FT   DISULFID    124    191
FT   DISULFID    156    170
FT   DISULFID    181    205
FT   SITE        179    179       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
FT   VARIANT      20     20       I -> V.
FT   STRAND       10     10
FT   STRAND       13     14
FT   TURN         17     18
FT   TURN         21     22
FT   STRAND       23     27
FT   STRAND       31     39
FT   TURN         40     41
FT   STRAND       42     45
FT   HELIX        47     49
FT   STRAND       55     58
FT   STRAND       62     62
FT   TURN         63     64
FT   STRAND       71     80
FT   TURN         82     83
FT   TURN         86     88
FT   TURN         90     91
FT   STRAND       94     98
FT   STRAND      112    112
FT   TURN        120    121
FT   STRAND      123    128
FT   STRAND      142    142
FT   STRAND      144    150
FT   HELIX       153    159
FT   TURN        161    163
FT   TURN        166    167
FT   STRAND      168    171
FT   TURN        174    175
FT   STRAND      179    179
FT   TURN        182    183
FT   TURN        185    186
FT   STRAND      188    191
FT   TURN        192    193
FT   STRAND      194    202
FT   STRAND      207    207
FT   TURN        208    209
FT   STRAND      210    210
FT   STRAND      212    216
FT   HELIX       217    219
FT   HELIX       221    230
SQ   SEQUENCE   231 AA;  24409 MW;  A0A125CF7FC138C2 CRC64;


  Query Match           61.6%;  Score 85;  DB 1;  Length 231;
  Best Local Similarity  52.0%;  Pred. No. 2.6e-06;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      27 NSGSHFCGGSLINSQWVVSAAHCYK 51


RESULT   5
TRY3_CHICK
ID   TRY3_CHICK     STANDARD;      PRT;   248 AA.
AC   Q90629;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN II-P29 PRECURSOR (EC 3.4.21.4).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=95251611; PubMed=7733885;
RA   Wang K., Gan L., Lee I., Hood L.E.;
RT   "Isolation and characterization of the chicken trypsinogen gene
RT   family.";
RL   Biochem. J. 307:471-479(1995).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE
CC       LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U15157; AAA79914.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL        1     16       BY SIMILARITY.
```

```
FT   ACT_SITE    219    219       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     54     70       BY SIMILARITY.
FT   DISULFID    150    225       BY SIMILARITY.
FT   DISULFID    183    206       BY SIMILARITY.
FT   DISULFID    215    243       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    228    228       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   270 AA;  30166 MW;  1BE102DB86943401 CRC64;


  Query Match           61.6%;  Score 85;  DB 1;  Length 270;
  Best Local Similarity  70.0%;  Pred. No. 3e-06;
  Matches   14;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 RHFCGGALIHARFVMTAASC 23
           |||||:||| ::|:||| |
Db      51 RHFCGGSLIHPQWVLTAAHC 70


RESULT   8
TRY2_SALSA
ID   TRY2_SALSA     STANDARD;      PRT;   231 AA.
AC   P35032;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN II PRECURSOR (EC 3.4.21.4) (FRAGMENT).
OS   Salmo salar (Atlantic salmon).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
OX   NCBI_TaxID=8030;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=96035908; PubMed=7556223;
RA   Male R., Lorens J.B., Smals A.O., Torrissen K.R.;
RT   "Molecular cloning and characterization of anionic and cationic
RT   variants of trypsin from Atlantic salmon.";
RL   Eur. J. Biochem. 232:677-685(1995).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X70073; CAA49678.1; -.
DR   PIR; S31778; S31778.
DR   HSSP; P35031; 1BIT.
DR   MEROPS; S01.258; -.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   Multigene family.
FT   NON_TER      1      1
FT   SIGNAL      <1      4       POTENTIAL.
FT   PROPEP       5      9       ACTIVATION PEPTIDE.
FT   CHAIN       10    231       TRYPSIN II.
FT   ACT_SITE    49     49       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    93     93       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   185    185       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    16    145       BY SIMILARITY.
FT   DISULFID    34     50       BY SIMILARITY.
FT   DISULFID   118    218       BY SIMILARITY.
FT   DISULFID   125    191       BY SIMILARITY.
FT   DISULFID   156    170       BY SIMILARITY.
FT   DISULFID   181    205       BY SIMILARITY.
FT   SITE       179    179       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   231 AA;  24823 MW;  C54A1CAFE74FAE1B CRC64;


  Query Match           60.9%;  Score 84;  DB 1;  Length 231;
  Best Local Similarity  52.0%;  Pred. No. 3.7e-06;
  Matches   13;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:|::  :|::|| |:|
Db      28 NSGYHFCGGSLVNENWVVSAAHCYQ 52


RESULT   9
TRY1_BOVIN
ID   TRY1_BOVIN     STANDARD;      PRT;   243 AA.
AC   P00760;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   15-JUL-1999 (Rel. 38, Last annotation update)
DE   TRYPSINOGEN, CATIONIC PRECURSOR (EC 3.4.21.4) (BETA-TRYPSIN)
DE   (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RA   Okajima T., Maniwa M., Nagao S., Fujikawa H., Goto S.;
RL   Submitted (OCT-1994) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 15-243, AND DISULFIDE BONDS.
RX   MEDLINE=67168848; PubMed=5967094;
RA   Mikes O., Holeysovsky V., Tomasek V., Sorm F.;
RT   "Covalent structure of bovine trypsinogen. The position of the
RT   remaining amides.";
RL   Biochem. Biophys. Res. Commun. 24:346-352(1966).
RN   [3]
RP   REVISIONS.
RX   MEDLINE=72035053; PubMed=4399051;
RA   Hartley B.S.;
RT   "Homologies in serine proteinases.";
RL   Philos. Trans. R. Soc. Lond., B, Biol. Sci. 257:77-87(1970).
RN   [4]
RP   REVISIONS.
RX   MEDLINE=75146445; PubMed=1092332;
RA   Titani K., Ericsson L.H., Neurath H., Walsh K.A.;
RT   "Amino acid sequence of dogfish trypsin.";
RL   Biochemistry 14:1358-1366(1975).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF CALCIUM-BINDING SITE.
RX   MEDLINE=76072097; PubMed=512;
RA   Bode W., Schwager P.;
RT   "The refined crystal structure of bovine beta-trypsin at 1.8-A
RT   resolution. II. Crystallographic refinement, calcium binding site,
RT   benzamidine binding site and active site at pH 7.0.";
RL   J. Mol. Biol. 98:693-717(1975).
RN   [6]
RP   X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX   MEDLINE=77112431; PubMed=556951;
RA   Kossiakoff A.A., Chambers J.L., Kay L.M., Stroud R.M.;
RT   "Structure of bovine trypsinogen at 1.9-A resolution.";
RL   Biochemistry 16:654-664(1977).
RN   [7]
RP   DISULFIDE BONDS.
```

```
FT   TURN        178    179
FT   STRAND      180    183

  Query Match             60.9%;  Score 84;  DB 1;  Length 243;
  Best Local Similarity   52.0%;  Pred. No. 3.9e-06;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||:::::::|| |::
Db      39 NSGYHFCGGSLINSQWVVSAAHCYK 63


RESULT  10
TRY1_CANFA
ID   TRY1_CANFA     STANDARD;      PRT;   246 AA.
AC   P06871;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN, CATIONIC PRECURSOR (EC 3.4.21.4).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86284628; PubMed=3841794;
RA   Pinsky S.D., Laforge K.S., Scheele G.;
RT   "Differential regulation of trypsinogen mRNA translation: full-length
RT   mRNA sequences encoding two oppositely charged trypsinogen isoenzymes
RT   in the dog pancreas.";
RL   Mol. Cell. Biol. 5:2669-2676(1985).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11590; AAA30900.1; -.
DR   PIR; B26273; TRDGC.
DR   HSSP; P00761; 1EPT.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL        1     15
FT   PROPEP       16     23       ACTIVATION PEPTIDE.
FT   CHAIN        24    246       TRYPSIN, CATIONIC.
FT   ACT_SITE     63     63       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    107    107       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    200    200       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     30    160       BY SIMILARITY.
FT   DISULFID     48     64       BY SIMILARITY.
FT   DISULFID    132    233       BY SIMILARITY.
FT   DISULFID    139    206       BY SIMILARITY.
FT   DISULFID    171    185       BY SIMILARITY.
FT   DISULFID    196    220       BY SIMILARITY.
FT   SITE        194    194       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   246 AA;  26170 MW;  E9E5A1DE2391BBBB CRC64;
```

```
  Query Match             60.9%;  Score 84;  DB 1;  Length 246;
  Best Local Similarity   52.0%;  Pred. No. 3.9e-06;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||:::::::|| |::
Db      42 NSGYHFCGGSLINSQWVVSAAHCYK 66


RESULT  11
TRY2_BOVIN
ID   TRY2_BOVIN     STANDARD;      PRT;   247 AA.
AC   Q29463;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN, ANIONIC PRECURSOR (EC 3.4.21.4).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIESIAN HOLSTEIN; TISSUE=Pancreas;
RX   MEDLINE=91065383; PubMed=1701147;
RA   le Huerou I., Wicker C., Guilloteau P., Toullec R., Puigserver A.;
RT   "Isolation and nucleotide sequence of cDNA clone for bovine
RT   pancreatic anionic trypsinogen. Structural identity within the
RT   trypsin family.";
RL   Eur. J. Biochem. 193:767-773(1990).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54703; CAA38513.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.258; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal.
FT   SIGNAL        1     15       POTENTIAL.
FT   PROPEP       16     23       ACTIVATION PEPTIDE.
FT   CHAIN        24    247       TRYPSIN, ANIONIC.
FT   ACT_SITE     63     63       CHARGE RELAY SYSTEM.
FT   ACT_SITE    107    107       CHARGE RELAY SYSTEM.
FT   ACT_SITE    200    200       CHARGE RELAY SYSTEM.
FT   DISULFID     30    160       BY SIMILARITY.
FT   DISULFID     48     64       BY SIMILARITY.
FT   DISULFID    132    233       BY SIMILARITY.
FT   DISULFID    139    206       BY SIMILARITY.
FT   DISULFID    171    185       BY SIMILARITY.
FT   DISULFID    196    220       BY SIMILARITY.
FT   SITE        194    194       REQUIRED FOR SPECIFICITY.
SQ   SEQUENCE   247 AA;  26289 MW;  50A070495A7731DB CRC64;
```

```
          | | |||||:||:::::|::|| |::
Db      43 NAGYHFCGGSLINSQWVVSAAHCYK 67


RESULT 14
TRY1_CHICK
ID   TRY1_CHICK     STANDARD;      PRT;   248 AA.
AC   Q90627;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN I-P1 PRECURSOR (EC 3.4.21.4).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=95251611; PubMed=7733885;
RA   Wang K., Gan L., Lee I., Hood L.E.;
RT   "Isolation and characterization of the chicken trypsinogen gene
RT   family.";
RL   Biochem. J. 307:471-479(1995).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE
CC       LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U15155; AAA79912.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.151; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL       1     15       BY SIMILARITY.
FT   PROPEP      16     25       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN       26    248       TRYPSIN I-P1.
FT   ACT_SITE    65     65       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   109    109       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   202    202       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    32    162       BY SIMILARITY.
FT   DISULFID    50     66       BY SIMILARITY.
FT   DISULFID   134    235       BY SIMILARITY.
FT   DISULFID   141    208       BY SIMILARITY.
FT   DISULFID   173    187       BY SIMILARITY.
FT   DISULFID   198    222       BY SIMILARITY.
FT   SITE       196    196       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   248 AA;  26069 MW;  C4CF589912B23D98 CRC64;


  Query Match           60.1%;  Score 83;  DB 1;  Length 248;
  Best Local Similarity  52.0%;  Pred. No. 5.7e-06;
  Matches   13;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:|| :::|::|| |::
Db      44 NSGYHFCGGSLISSQWVLSAAHCYK 68


RESULT 15
TRY2_CHICK
ID   TRY2_CHICK     STANDARD;      PRT;   248 AA.
AC   Q90628;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   TRYPSIN I-P38 PRECURSOR (EC 3.4.21.4).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Pancreas;
RX   MEDLINE=95251611; PubMed=7733885;
RA   Wang K., Gan L., Lee I., Hood L.E.;
RT   "Isolation and characterization of the chicken trypsinogen gene
RT   family.";
RL   Biochem. J. 307:471-479(1995).
CC   -!- CATALYTIC ACTIVITY: PREFERENTIAL CLEAVAGE: ARG-, LYS-.
CC   -!- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC   -!- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE PANCREAS WHILE
CC       LOWER LEVELS ARE FOUND IN THE LIVER, SPLEEN AND THYMUS.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC       TRYPSIN FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U15156; AAA79913.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.258; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Digestion; Pancreas; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL       1     15       BY SIMILARITY.
FT   PROPEP      16     25       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN       26    248       TRYPSIN I-P38.
FT   ACT_SITE    65     65       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   109    109       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE   202    202       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID    32    162       BY SIMILARITY.
FT   DISULFID    50     66       BY SIMILARITY.
FT   DISULFID   134    235       BY SIMILARITY.
FT   DISULFID   141    208       BY SIMILARITY.
FT   DISULFID   173    187       BY SIMILARITY.
FT   DISULFID   198    222       BY SIMILARITY.
FT   SITE       196    196       REQUIRED FOR SPECIFICITY (BY SIMILARITY).
SQ   SEQUENCE   248 AA;  26087 MW;  78B79DD6FE15F0CE CRC64;


  Query Match           60.1%;  Score 83;  DB 1;  Length 248;
  Best Local Similarity  52.0%;  Pred. No. 5.7e-06;
  Matches   13;  Conservative    8;  Mismatches    4;  Indels    0;  Gaps    0;
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 5, 2002, 14:58:16 ; Search time 22.81 Seconds
(without alignments)
160.316 Million cell updates/sec

Title: US-09-657-276-1032
Perfect score: 138
Sequence: 1 NQGRHFCGGALIHARFVMTAASCFQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	98	71.0	254	11	O08809	O08809 mus musculu
2	98	71.0	254	11	Q61096	Q61096 mus musculu
3	98	71.0	258	6	Q9GME1	Q9gme1 ornithorhyn
4	93	67.4	200	4	Q9UQD8	Q9uqd8 homo sapien
5	92	66.7	799	11	Q9DBI0	Q9dbi0 mus musculu
6	88	63.8	275	6	Q9N2D1	Q9n2d1 sus scrofa
7	85	61.6	117	13	Q9DFJ6	Q9dfj6 gillichthys
8	85	61.6	247	13	O42608	O42608 petromyzon
9	85	61.6	247	13	O42158	O42158 petromyzon
10	84	60.9	247	11	Q9D7Y7	Q9d7y7 mus musculu
11	84	60.9	247	11	Q9CPN9	Q9cpn9 mus musculu
12	83	60.1	269	5	Q9V929	Q9v929 drosophila
13	83	60.1	567	4	Q13208	Q13208 homo sapien
14	83	60.1	648	4	Q9H1V4	Q9h1v4 homo sapien
15	83	60.1	711	4	Q14870	Q14870 homo sapien
16	82	59.4	166	4	Q9UQI8	Q9uqi8 homo sapien
17	82	59.4	244	13	O42159	O42159 petromyzon
18	82	59.4	245	5	Q9XY47	Q9xy47 ctenocephal
19	82	59.4	245	13	O42160	O42160 petromyzon
20	82	59.4	290	4	Q9BQR3	Q9bqr3 homo sapien
21	81	58.7	190	4	Q9H2Y6	Q9h2y6 homo sapien
22	81	58.7	235	4	Q9BZJ3	Q9bzj3 homo sapien
23	81	58.7	242	4	O95824	O95824 homo sapien
24	81	58.7	246	11	Q9Z1R9	Q9z1r9 mus musculu
25	81	58.7	246	11	Q9QUK9	Q9quk9 mus musculu
26	81	58.7	246	11	Q9R0T7	Q9r0t7 mus musculu
27	81	58.7	247	4	Q9UQV3	Q9uqv3 homo sapien
28	81	58.7	259	4	Q15665	Q15665 homo sapien
29	81	58.7	268	5	Q9XY56	Q9xy56 ctenocephal
30	81	58.7	284	5	O96089	O96089 haemaphysal
31	81	58.7	355	5	Q9NFU1	Q9nfu1 anopheles g
32	81	58.7	542	5	Q9I7J3	Q9i7j3 drosophila
33	80	58.0	178	13	O93594	O93594 dicentrarch
34	80	58.0	213	4	Q9H2Y4	Q9h2y4 homo sapien
35	80	58.0	233	4	O95827	O95827 homo sapien
36	80	58.0	237	13	Q91515	Q91515 fugu rubrip
37	80	58.0	238	13	Q9W7Q6	Q9w7q6 paralichthy
38	80	58.0	240	13	Q98TH0	Q98th0 engraulis j
39	80	58.0	242	13	Q92099	Q92099 paranotothe
40	80	58.0	242	13	O93266	O93266 pleuronecte
41	80	58.0	242	13	Q9W7Q7	Q9w7q7 paralichthy
42	80	58.0	247	13	Q9W7Q5	Q9w7q5 paralichthy
43	80	58.0	251	4	Q9H2Y5	Q9h2y5 homo sapien
44	80	58.0	258	5	Q9W5U8	Q9w5u8 drosophila
45	80	58.0	267	4	Q15664	Q15664 homo sapien

ALIGNMENTS

```
RESULT   1
O08809
ID   O08809      PRELIMINARY;      PRT;   254 AA.
AC   O08809;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   PROTEINASE 3 (PRE-PRO-PROTEINASE 3).
GN   PRTN3.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129SV;
RA   Jenne D.E., Froehlich L., Hummel A.M., Specks U.;
RL   FEBS Lett. 0:0-0(0).
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; U97073; AAB58055.1; -.
DR   HSSP; P24158; 1FUJ.
DR   MGD; MGI:893580; Prtn3.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   254 AA;  27610 MW;  0E74B989A71179CA CRC64;


  Query Match            71.0%;  Score 98;  DB 11;  Length 254;
  Best Local Similarity  73.9%;  Pred. No. 4.7e-08;
  Matches   17;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       3 GRHFCGGALIHARFVMTAASCFQ 25
           | ||||| ||| |||:||| | |
Db      54 GSHFCGGTLIHPRFVLTAAHCLQ 76
```

```
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   1300008A22RIK PROTEIN.
GN   1300008A22RIK.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=LIVER;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
CC   -!- SIMILARITY: TO LOW DENSITY LIPOPROTEIN (LDL) RECEPTOR CLASS A
CC       (LDLRA) DOMAIN.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AK004939; BAB23684.1; -.
DR   MGD; MGI:1919003; 1300008A22Rik.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000859; CUB.
DR   InterPro; IPR002172; LDL_recept_A.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00057; ldl_recept_a; 3.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00261; LDLRECEPTOR.
DR   SMART; SM00042; CUB; 1.
DR   SMART; SM00192; LDLa; 3.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS01180; CUB; 1.
DR   PROSITE; PS50068; LDLRA_2; 3.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Glycoprotein; Hydrolase; Serine protease.
SQ   SEQUENCE   799 AA;  89557 MW;  16315A646A4D5288 CRC64;


  Query Match             66.7%;  Score 92;  DB 11;  Length 799;
  Best Local Similarity   70.8%;  Pred. No. 1.3e-06;
  Matches   17;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCFQ 25
           :||| ||||||  |:|:||| |||
Db     585 RGRHICGGALIADRWVITAAHCFQ 608


RESULT   6
Q9N2D1
ID   Q9N2D1      PRELIMINARY;      PRT;   275 AA.
AC   Q9N2D1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MINIATURE SWINE MAST CELL TRYPTASE PRECURSOR.
GN   PMCT7.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Iida K., Yamashita M., Kido H.;
RL   Submitted (FEB-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LUNG;
RA   Chen Y., Shiota M., Ohuchi M., Towatari T., Tashiro J., Murakami M.,
RA   Yano M., Yang B., Kido H.;
RT   "A novel trypsin-type protease from pig lungs, triggering infection by
RT   peumotropic viruses: Purification and characterizaion.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AB038652; BAA93614.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    275       TRYPTASE.
SQ   SEQUENCE   275 AA;  30439 MW;  ACC582647FCCB973 CRC64;


  Query Match             63.8%;  Score 88;  DB 6;  Length 275;
  Best Local Similarity   66.7%;  Pred. No. 2.1e-06;
  Matches   14;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 RHFCGGALIHARFVMTAASCF 24
           :|||||:||| ::|:||| ||
Db      56 KHFCGGSLIHPQWVLTAAHCF 76


RESULT   7
Q9DFJ6
ID   Q9DFJ6      PRELIMINARY;      PRT;   117 AA.
AC   Q9DFJ6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   TRYPSINOGEN 2 PRECURSOR (FRAGMENT).
OS   Gillichthys seta.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Gobioidei;
OC   Gobiidae; Gillichthys.
OX   NCBI_TaxID=79683;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LIVER;
RA   Gracey A.Y., Troll J.V., Somero G.N.;
RT   "Hypoxia regulated gene expression.";
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AF266240; AAG13359.1; -.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Hydrolase; Serine protease.
```

```
DR   MGD; MGI:1914623; 2210010C04Rik.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   247 AA;  26407 MW;  84D474DB322B3A55 CRC64;


  Query Match             60.9%;  Score 84;  DB 11;  Length 247;
  Best Local Similarity   52.0%;  Pred. No. 8.5e-06;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      43 NSGYHFCGGSLINSQWVVSAAHCYK 67


RESULT  11
Q9CPN9
ID   Q9CPN9      PRELIMINARY;      PRT;   247 AA.
AC   Q9CPN9;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   2210010C04RIK PROTEIN (TRYPSINOGEN 7).
GN   2210010C04RIK OR TRYPSINOGEN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE, AND PANCREAS;
RX   MEDLINE=21085660; PubMed=11217851;
RA   Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA   Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA   Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA   Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA   Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA   Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA   Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA   Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA   Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA   Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA   Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA   Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA   Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA   Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA   Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA   Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA   Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA   Hayashizaki Y.;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/C;
RX   MEDLINE=21103195; PubMed=11160223;
RA   Chen F., Rowen L., Hood L., Rothenberg E.V.;
RT   "Differential transcriptional regulation of individual TCR Vbeta
RT   segments before gene rearrangement.";
RL   J. Immunol. 166:1771-1780(2001).
CC   -!- SIMILARITY: TO SERINE PROTEASES, TRYPSIN FAMILY.
CC   -!- SIMILARITY: TO CHYMOTRYPSIN SERINE PROTEASE FAMILY (S1).
DR   EMBL; AK008460; BAB25680.1; -.
DR   EMBL; AK007773; BAB25246.1; -.
DR   EMBL; AE000663; AAB69044.1; -.
DR   MGD; MGI:1914623; 2210010C04Rik.
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
KW   Hydrolase; Serine protease.
SQ   SEQUENCE   247 AA;  26422 MW;  B8C5767B182D9AAD CRC64;


  Query Match             60.9%;  Score 84;  DB 11;  Length 247;
  Best Local Similarity   52.0%;  Pred. No. 8.5e-06;
  Matches   13;  Conservative    9;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           | | |||||:||::::|::|| |::
Db      43 NSGYHFCGGSLINSQWVVSAAHCYK 67


RESULT  12
Q9V929
ID   Q9V929      PRELIMINARY;      PRT;   269 AA.
AC   Q9V929;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   CG11192 PROTEIN.
GN   CG11192.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
```

```
DR   InterPro; IPR001314; Chymotrypsin.
DR   InterPro; IPR000001; Kringle.
DR   InterPro; IPR003014; PAN.
DR   InterPro; IPR003609; Pan_app.
DR   InterPro; IPR001254; Trypsin.
DR   Pfam; PF00051; kringle; 4.
DR   Pfam; PF00024; PAN; 1.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00018; KRINGLE.
DR   SMART; SM00130; KR; 4.
DR   SMART; SM00473; PAN_AP; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS00021; KRINGLE_1; 4.
DR   PROSITE; PS50070; KRINGLE_2; 4.
KW   Hydrolase; Serine protease; Signal.
FT   SIGNAL        1     18       POTENTIAL.
FT   CHAIN        19    483       MACROPHAGE-STIMULATING PROTEIN.
FT   CHAIN       484    711       MACROPHAGE-STIMULATING PROTEIN.
SQ   SEQUENCE   711 AA;  80353 MW;  24E136DDE0A31F66 CRC64;


  Query Match             60.1%;  Score 83;  DB 4;  Length 711;
  Best Local Similarity   52.2%;  Pred. No. 3.4e-05;
  Matches   12;  Conservative    7;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 QGRHFCGGALIHARFVMTAASCF 24
           ||:|||||:|:  ::::||  ||
Db     502 QGQHFCGGSLVKEQWILTAROCF 524


Search completed: April  5, 2002, 15:00:56
Job time: 160 sec
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: April 5, 2002, 14:56:16 ; Search time 12.45 Seconds
(without alignments)
45.187 Million cell updates/sec

Title: US-09-657-276-1032
Perfect score: 138
Sequence: 1 NQGRHFCGGALIHARFVMTAASCFQ 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	138	100.0	25	1	US-07-969-931-8	Sequence 8, Appli
2	138	100.0	25	1	US-07-855-417A-8	Sequence 8, Appli
3	138	100.0	25	1	US-08-235-399-1	Sequence 1, Appli
4	138	100.0	25	1	US-08-482-328-1	Sequence 1, Appli
5	138	100.0	25	2	US-08-840-519-1	Sequence 1, Appli
6	138	100.0	25	3	US-09-260-373-1	Sequence 1, Appli
7	138	100.0	25	3	US-09-258-934-1	Sequence 1, Appli
8	138	100.0	221	2	US-08-925-708-1	Sequence 1, Appli
9	138	100.0	222	1	US-07-969-931-9	Sequence 9, Appli
10	138	100.0	222	1	US-07-855-417A-9	Sequence 9, Appli
11	138	100.0	222	2	US-08-491-204A-18	Sequence 18, Appl
12	138	100.0	225	4	US-08-944-483-32	Sequence 32, Appl
13	137	99.3	25	3	US-09-258-934-38	Sequence 38, Appl
14	135	97.8	25	3	US-09-258-934-36	Sequence 36, Appl
15	135	97.8	25	3	US-09-258-934-37	Sequence 37, Appl
16	135	97.8	25	3	US-09-258-934-40	Sequence 40, Appl
17	135	97.8	25	3	US-09-258-934-41	Sequence 41, Appl
18	135	97.8	25	3	US-09-258-934-42	Sequence 42, Appl
19	134	97.1	25	3	US-09-258-934-11	Sequence 11, Appl
20	134	97.1	25	3	US-09-258-934-39	Sequence 39, Appl
21	133	96.4	25	3	US-09-258-934-9	Sequence 9, Appli
22	133	96.4	25	3	US-09-258-934-10	Sequence 10, Appl
23	132	95.7	25	3	US-09-258-934-27	Sequence 27, Appl
24	132	95.7	25	3	US-09-258-934-28	Sequence 28, Appl
25	128	92.8	25	3	US-09-258-934-3	Sequence 3, Appli
26	128	92.8	25	3	US-09-258-934-5	Sequence 5, Appli
27	126	91.3	25	3	US-09-258-934-29	Sequence 29, Appl
28	124	89.9	25	3	US-09-258-934-17	Sequence 17, Appl
29	124	89.9	25	3	US-09-258-934-20	Sequence 20, Appl
30	123	89.1	25	3	US-09-258-934-15	Sequence 15, Appl
31	123	89.1	25	3	US-09-258-934-16	Sequence 16, Appl
32	123	89.1	25	3	US-09-258-934-18	Sequence 18, Appl
33	123	89.1	25	3	US-09-258-934-19	Sequence 19, Appl
34	122	88.4	25	3	US-09-258-934-30	Sequence 30, Appl
35	122	88.4	25	3	US-09-258-934-31	Sequence 31, Appl
36	122	88.4	25	3	US-09-258-934-33	Sequence 33, Appl
37	122	88.4	25	3	US-09-258-934-34	Sequence 34, Appl
38	118	85.5	25	3	US-09-258-934-7	Sequence 7, Appli
39	116	84.1	25	3	US-09-258-934-32	Sequence 32, Appl
40	116	84.1	25	3	US-09-258-934-35	Sequence 35, Appl
41	110	79.7	20	1	US-07-969-931-7	Sequence 7, Appli
42	110	79.7	20	1	US-07-855-417A-7	Sequence 7, Appli
43	110	79.7	20	1	US-08-235-399-2	Sequence 2, Appli
44	110	79.7	20	1	US-08-482-328-2	Sequence 2, Appli
45	110	79.7	20	2	US-08-840-519-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-07-969-931-8
; Sequence 8, Application US/07969931
; Patent No. 5458874
; GENERAL INFORMATION:
;   APPLICANT: Pereira, Heloise Anne
;   APPLICANT: Spitznagel, John K.
;   TITLE OF INVENTION: Chemotactic, Antibiotic and
;   TITLE OF INVENTION: Lipopolysaccharide-Binding Peptide Fragments of CAP37
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Needle & Rosenberg, P.C.
;     STREET: 133 Carnegie Way N.W., Suite 400
;     CITY: Atlanta
;     STATE: Georgia
;     COUNTRY: USA
;     ZIP: 30303
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/07/969,931
;     FILING DATE: 19921030
;     CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/543,151
;     FILING DATE: 25-JUN-1990
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 07/375,739
;     FILING DATE: 05-JUL-1989
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Severson, Mary L.
;     REGISTRATION NUMBER: 34,927
;     REFERENCE/DOCKET NUMBER: 0510.024
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (404) 688-0770
;     TELEFAX: (404) 688-9880
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 25 amino acids
;     TYPE: AMINO ACID
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
;   FRAGMENT TYPE: internal
;   ORIGINAL SOURCE: 20-44aa of mature CAP37 protein
US-07-969-931-8
```

```
;      STREET:  9400 No. 5627262th Broadway, Suite 420
;      CITY:  Oklahoma City
;      STATE:  Oklahoma
;      COUNTRY:  United States of America
;      ZIP:  73114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5, 1.44 Mb High Density Diskette
;      COMPUTER:  IBM AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.00
;      SOFTWARE:  WordPerfect 5.1 (saved in ASCII format)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,328
;      FILING DATE:  Filed Herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/235,399
;      FILING DATE:  April 29, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,931
;      FILING DATE:  October 30, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/855,417
;      FILING DATE:  March 18, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/543,151
;      FILING DATE:  June 25, 1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/375,739
;      FILING DATE:  July 5, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corbett, Christopher W., Ph.D.
;      REGISTRATION NUMBER:  36,109
;      REFERENCE/DOCKET NUMBER:  5820.360
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  405-478-5344
;      TELEFAX:  405-478-5349
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:  20-44aa of mature CAP37 protein
US-08-482-328-1


  Query Match           100.0%;  Score 138;  DB 1;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   5
US-08-840-519-1
; Sequence 1, Application US/08840519
; Patent No. 5877151
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne
;    TITLE OF INVENTION:  Method and Composition for The Treatment
;    TITLE OF INVENTION:  of Septic Shock
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christopher W. Corbett, Ph.D.
;      ADDRESSEE:  Dunlap & Codding, P.C.
;      STREET:  9400 No. 5877151th Broadway, Suite 420
;      CITY:  Oklahoma City
;      STATE:  Oklahoma
;      COUNTRY:  United States of America
;      ZIP:  73114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5, 1.44 Mb High Density Diskette
;      COMPUTER:  IBM AT Compatible
;      OPERATING SYSTEM:  MS-DOS 5.00
;      SOFTWARE:  WordPerfect 5.1 (saved in ASCII format)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/840,519
;      FILING DATE:  21-APR-1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/482,328
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  08/235,399
;      FILING DATE:  April 29, 1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/969,931
;      FILING DATE:  October 30, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/855,417
;      FILING DATE:  March 18, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/543,151
;      FILING DATE:  June 25, 1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/375,739
;      FILING DATE:  July 5, 1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Corbett, Christopher W., Ph.D.
;      REGISTRATION NUMBER:  36,109
;      REFERENCE/DOCKET NUMBER:  5820.360
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  405-478-5344
;      TELEFAX:  405-478-5349
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
;    FRAGMENT TYPE:  internal
;    ORIGINAL SOURCE:  20-44aa of mature CAP37 protein
US-08-840-519-1


  Query Match           100.0%;  Score 138;  DB 2;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.3e-14;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db      1 NQGRHFCGGALIHARFVMTAASCFQ 25


RESULT   6
US-09-260-373-1
; Sequence 1, Application US/09260373
; Patent No. 6071879
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne
;    TITLE OF INVENTION:  Method and Composition for The Treatment
;    TITLE OF INVENTION:  of Septic Shock
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Christopher W. Corbett, Ph.D.
;      ADDRESSEE:  Dunlap & Codding, P.C.
;      STREET:  9400 No. 6071879th Broadway, Suite 420
;      CITY:  Oklahoma City
;      STATE:  Oklahoma
;      COUNTRY:  United States of America
;      ZIP:  73114
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5, 1.44 Mb High Density Diskette
;      COMPUTER:  IBM AT Compatible
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/969,931
;      FILING DATE:  19921030
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/543,151
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/375,739
;      FILING DATE:  05-JUL-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Severson, Mary L.
;      REGISTRATION NUMBER:  34,927
;      REFERENCE/DOCKET NUMBER:  0510.024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  222 amino acids
;      TYPE:  AMINO ACID
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-07-969-931-9


  Query Match           100.0%;  Score 138;  DB 1;  Length 222;
  Best Local Similarity  100.0%;  Pred. No. 6.9e-13;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db     20 NQGRHFCGGALIHARFVMTAASCFQ 44


RESULT  10
US-07-855-417A-9
; Sequence 9, Application US/07855417A
; Patent No. 5484885
;  GENERAL INFORMATION:
;    APPLICANT:  Pereira, Heloise Anne
;    APPLICANT:  Spitznagel, John K.
;    TITLE OF INVENTION:  Chemotactic, Antibiotic and
;    TITLE OF INVENTION:  Lipopolysaccharide-Binding Peptide Fragments of CAP37
;    NUMBER OF SEQUENCES:  29
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Needle & Rosenberg, P.C.
;      STREET:  133 Carnegie Way N.W., Suite 400
;      CITY:  Atlanta
;      STATE:  Georgia
;      COUNTRY:  USA
;      ZIP:  30303
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/855,417A
;      FILING DATE:  18-MAR-1992
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/543,151
;      FILING DATE:  25-JUN-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/375,739
;      FILING DATE:  05-JUL-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Severson, Mary L.
;      REGISTRATION NUMBER:  34,927
;      REFERENCE/DOCKET NUMBER:  0510.024
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (404) 688-0770
;      TELEFAX:  (404) 688-9880
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  222 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FRAGMENT TYPE:  internal
US-07-855-417A-9


  Query Match           100.0%;  Score 138;  DB 1;  Length 222;
  Best Local Similarity  100.0%;  Pred. No. 6.9e-13;
  Matches   25;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 NQGRHFCGGALIHARFVMTAASCFQ 25
          |||||||||||||||||||||||||
Db     20 NQGRHFCGGALIHARFVMTAASCFQ 44


RESULT  11
US-08-491-204A-18
; Sequence 18, Application US/08491204A
; Patent No. 5837247
;  GENERAL INFORMATION:
;    APPLICANT:  Oppenheim, Joost B.
;    APPLICANT:  Chertov, Oleg
;    APPLICANT:  Michiel, Dennis F.
;    APPLICANT:  Xu, Luoling
;    APPLICANT:  Wang, Ji Ming
;    APPLICANT:  Murphy, William J.
;    APPLICANT:  Longo, Dan L.
;    APPLICANT:  Taub, Dennis D.
;    TITLE OF INVENTION:  CHEMOTACTIC AGENTS FOR T-CELLS
;    NUMBER OF SEQUENCES:  18
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genetics Institute, Inc.
;      STREET:  87 CambridgePark Drive
;      CITY:  Cambridge
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02140
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/491,204A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brown, Scott A.
;      REGISTRATION NUMBER:  32,724
;      REFERENCE/DOCKET NUMBER:  GI5260
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617) 498-8224
;      TELEFAX:  (617) 876-5851
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  222 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    HYPOTHETICAL:  NO
US-08-491-204A-18
```

US-09-258-934-37

Query Match 97.8%; Score 135; DB 3; Length 25;
Best Local Similarity 96.0%; Pred. No. 2.2e-13;
Matches 24; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 NQGRHFCGGALIHARFVMTAASCFQ 25
           |||||||||||||||:|||||||||
Db       1 NQGRHFCGGALIHARFLMTAASCFQ 25
```

Search completed: April 5, 2002, 15:00:06
Job time: 230 sec